ESD-TR-74-345

ADAPT Report 74-4
November 1974

FINAL REPORT

APPLICATION OF ADAPT TO QUICK LOOK CLASSIFICATION OF COMPOSITE RADAR SIGNATURES

BY

Herbert E. Hunter

Prepared By

ADAPT Service CORPORATION

READING, MASSACHUSETTS 01867

FOR

Massachusetts Institute of Technology
Lincoln Laboratory

Under

Purchase Order No. C-563

Prime Contract No. F19628-73-C-0002

This Report Covers the Period

November 1973 to November 1974

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Eugene C. Raabe

Eugene C. Raabe, Lt. Col., USAF
Chief, ESD Lincoln Laboratory Project Office

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ESD-TR-74-345	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Application of ADAPT to Quick Look Classification of Composite Radar Signatures		5. TYPE OF REPORT & PERIOD COVERED Final Report (November 1973 to November 1974) 6. PERFORMING ORG. REPORT NUMBER ADAPT 74-4
7. AUTHOR(s) Hunter, Herbert E.		8. CONTRACT OR GRANT NUMBER(s) F19628-73-C-0002
9. PERFORMING ORGANIZATION NAME AND ADDRESS ADAPT Service Corporation under Purchase Order No. C-563 to M.I.T. Lincoln Laboratory		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Project No. 7X263304D215
11. CONTROLLING OFFICE NAME AND ADDRESS Ballistic Missile Defense Program Office Department of the Army 1320 Wilson Boulevard Arlington, VA 22209		12. REPORT DATE November 1974 13. NUMBER OF PAGES 190
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Electronic Systems Division Hanscom AFB Bedford, MA 01731		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Discrimination	Classification
Radar	Pattern Recognition
Signature Analysis	ADAPT

20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The ADAPT Computer Programs (a series of empirical analysis programs based on the concept that pattern recognition and regression analysis should be preceded by a reduction in the dimensionality based on the Karhunen-Loeve expansion) were applied to composite radar signatures of two classes of targets to develop and study classification laws and mechanisms for separating the two classes. The analysis showed that the 10 pulse signatures were adequate for the classifications desired and illustrated several mechanisms for classifications based on the differences in the characteristics of the means of the two classes. The

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Box 20. ABSTRACT (Continued):

analysis also showed that the most effective classification scheme for separating the two types of targets was based on the significantly greater variation observed in the second class of targets.

ACKNOWLEDGMENT

The author would like to thank the many personnel at Lincoln Laboratory for their encouragement and many helpful discussions of the results which are presented in this report. Specifically, thanks are due to Mr. Allan Grometstein for his assistance in selecting the problem areas and to Dr. William Schoendorf for his many helpful discussions especially in understanding the data and the definition of the source of the phase structure and its role in the discrimination. Thanks are also due to Mrs. Patricia Gaudet of the ADAPT Service Corporation for performing much of the ADAPT Data Processing included in this report.

ABSTRACT

The ADAPT computer programs, (a series of empirical data analysis programs based on the concept that pattern recognition and regression analysis should be preceded by a reduction in the dimensionality based on the Karhunen-Loeve expansion), were applied to composite radar signatures of two classes of targets to develop, study and recommend classification laws and mechanisms for separating the two classes. The study was based on two different sets of data.The first data set consisted of the incoherent radar back scatter from 200 pulses and was made up of 225 learning cases plus 85 independent test samples. This data set was used for exploratory analysis to guide the processing required for the second data set and to determine the effect of varying from 10 to 200 pulses. The second data set consisted of the coherent radar back scatter from 10 radar pulses for 260 learning cases and 2,140 independent test cases. The first data set included seven different target types and the second data set included five different target types. The remainder of the variation in both of these data sets was due to variations created by simulations of target dynamics and radar-target geometry.

This report summarizes the results obtained from the entire study and presents the detailed results for the analysis of the second data set. The analysis of the first data set showed that the 10 pulse signatures were adequate for the classifications desired and illustrated several mechanisms for classifications based on the differences in the characteristics of the means of the two classes. The analysis of the second data set showed that the most effective classification scheme for separating the two types of targets was based on the significantly greater variation observed in the second class of targets. This classification mechanism could be most effectively utilized by using a 10 or less pulse radar signature constructed from the separate squares of the real component and the imaginary component of the principal polarization. Linear classification schemes considered proved superior to the non-linear schemes and the recommended procedure is a multi-step procedure consisting of successively thresholding the projections of the data vectors on the ADAPT optimal coordinates (i.e. the principle components of this data set). Using this classification procedure, the estimated performances is a detection probability of .99 with no observed false alarms in the data set used.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (CONT'D)

LIST OF TABLES (CONT'D)

LIST OF FIGURES

LIST OF FIGURES (CONT'D)

LIST OF FIGURES (CONT'D)

LIST OF FIGURES (CONT'D)

LIST OF FIGURES (CONT'D)

LIST OF FIGURES (CONT'D)

LIST OF FIGURES (CONT'D)

LIST OF FIGURES (CONT'D)

1.0 INTRODUCTION

This report presents the results of a study which employed the ADAPT techniques to derive and compare numerous quick-look classification schemes for utilizing composite radar signatures to derive techniques suitable for real time classification of two types of targets. The first type of target is designated as the 101 Class. The second type of target is designated as the 20X and the 21X Class. The 20X Class was made up of six sub classes designated as 201, 202, 204, 205, 206 and 207. The 21X Class was made up of four sub classes designated 211, 212, 213 and 214.

The study was divided into two tasks. The first task considered the development of classification laws to separate the 101 Class from the 20X Class using data which could be derived from incoherent radar signatures. The effect of the amount of radar data required to perform the classification was evaluated by considering 10, 30 and 200 pulse data sets. The major objectives of this first task were: 1) to determine the best approach to processing the data, 2) to provide an estimate of the potential for separating the two types of targets and 3) to determine the number of radar pulses that would be required to provide adequate information to accomplish the classification.

Task 2 of this study utilized coherent radar signatures. Based on the results of Task 1, the study was limited to data obtained from 10 radar pulses, but further investigation of radar resources was carried out by considering the effects of using only the principal polarization of the data and the effect of using only the data in the time domain. In addition to evaluating the effect of the radar resources the second task had as a major objective the development and demonstration of a classification procedure for separating these two classes of targets.

Section 2 of this report will present the summary of the results of the entire study obtained in both the Task 1 and the Task 2 studies. The detailed analysis leading to these results for the Task 1 studies is presented in Reference 1 and will not be repeated here. The detail analysis leading to these results for Task 2 is presented

in Sections 3 through 5 of this report.

These studies were performed using the empirical analysis techniques incorporated in the ADAPT programs which have been under development since 1964 and are described in considerable detail in Reference 1. The ADAPT approach is based on the concept that one should precede an empirical analysis by determining the optimal (in Karhunen-Loeve sense) representation of the data. This representation of the data may then be used instead of the original data in performing the empirical analysis. The optimal representation is obtained by transforming to the ordered optimal coordinate system defined by the Karhunen-Loeve expansion. This optimal representation is also known as the principle components expansion, as an expansion in the optimum empirical orthogonal functions or as an eigenfunction expansion. The specific techniques utilized to accomplish this are described in more detail in Reference 1. The ADAPT program can accomplish this transformation for an unlimited number of data vectors of up to 2,000 components each. The data are usually adequately represented for most analyses in this optimal coordinate system using one-tenth to one-hundredth as many components as in the original system. In addition, to savings in computation which are realized as a result of this reduction in dimensionality, one also achieves other significant advantages such as an analysis of the learning data, a decrease in the number of learning cases required and an empirical validity criteria. These advantages resulting from this optimal representation will enhance the ability to perform empirical analysis in general and the development of the classification algorithms required for the present study in particular.

The radar back scatter data used for Task 1 of this study were the amplitude and quantities which could be derived from the amplitude. Pre-processing was performed on the data by taking the Fourier transform of the data, by taking logarithm of the amplitudes involved and by an equalization of the data histories by a technique which is described in detail in Reference 1. The data histories utilized for the analysis initially included all these pre-processings and an ADAPT output, the relative importance vector, was examined to determine which of these pre-processings appeared to make the greater contribution to the decisions. Based on the results of the examination of these relative importance vectors, the final analysis

used 225 data histories made up of the tandem amalgamation of three sets consisting of the amplitude, amplitude and phase of the Fourier transform of the data derived from both the principal and orthogonal polarization of the radar back scatter. Since the logarithm was taken of the amplitude measurements and not of the phase measurements, the equalization process utilized highlighted different types of behavior for the amplitude measurements then for the phase measurements in the Task 1 studies.

The learning data utilized for the Task 2 studies consisted of the real and imaginary portions of the principal and orthogonal polarization of the coherent radar return from 260 targets. The initial studies combine this with the corresponding real and imaginary parts of the Fourier transform of this data. Equalization was not performed on this data set. However, in addition to evaluating the performance using this entire data set, the effect of using only the principal polarization portion of this data set, only the time domain portion of this data set and the effect of squaring each component of the data prior to the processing was also evaluated. For all of the data histories used in both Task 1 and Task 2, the final pre-processing step was to form a single average of all the data histories and to subtract this average from the individual history before processing through the ADAPT programs to find the optimal representation.

Although a number of different methods were used to estimate the performance of classification schemes during the development and selection of classification procedures, the performance of the recommended procedure was evaluated using the experimentally calculated Receiver Operating Curve (ROC) derived from 1800 independent test cases. These test cases were derived in the same manner and for the same targets as the Task 2 learning data but for different dynamic histories. Two sets of test data were used for the Task 2 analysis. The first consisted of 340 cases, supplied with the learning data. This test set will be called the 340 case test set and included 100-101 Class signals and 60 signals for each of the 21X target types. The second independent test set used for the Task 2 studies will be called the 1800 case test set and consisted of an additional 1800 cases. This test set included 1000-101 Class signals and 200 signals for each of the 21X target types.

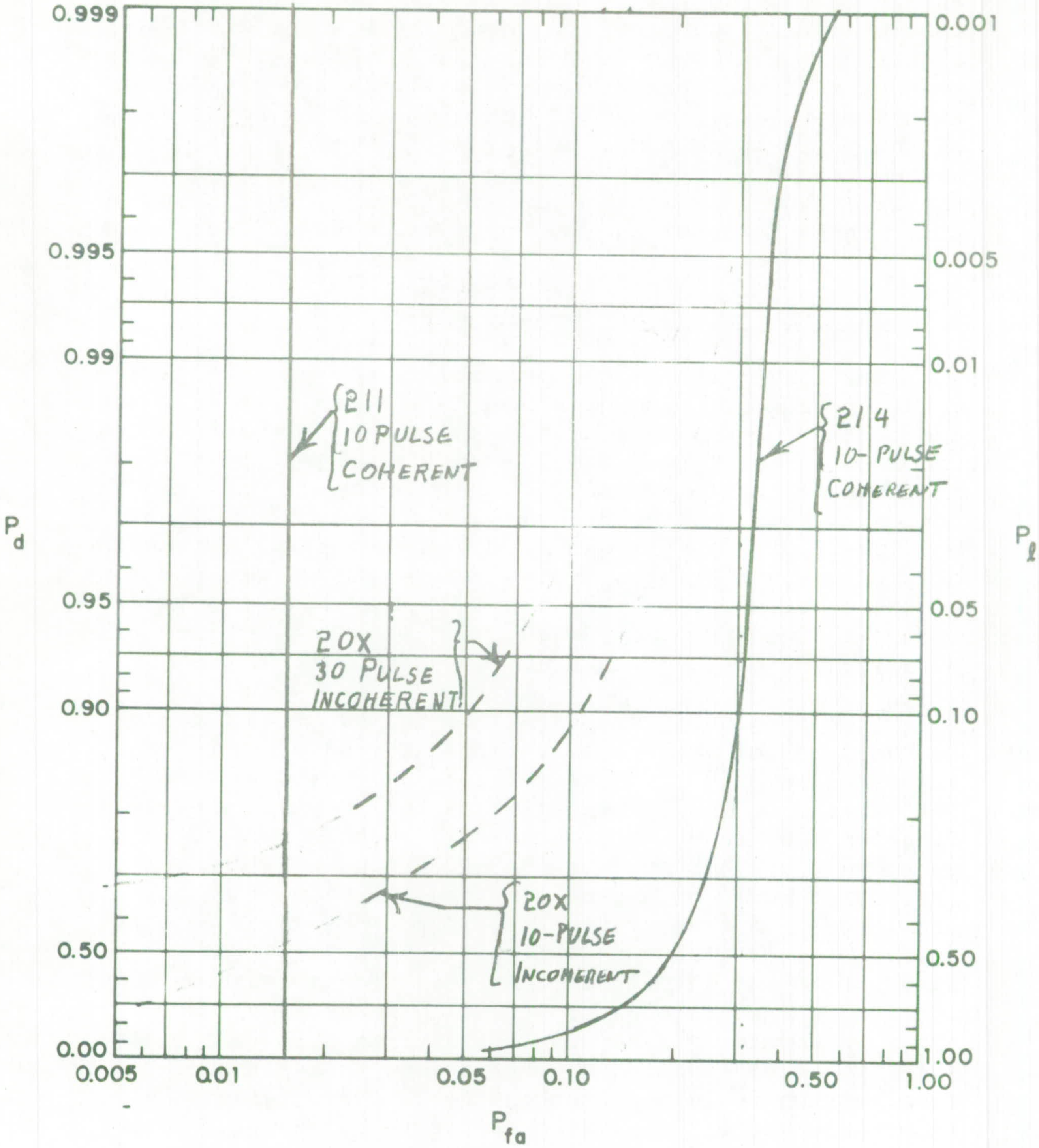
2.0 RESULTS AND RECOMMENDATIONS

The results and recommendations which will be presented in this section of the report are based on the analysis described in detail in Sections 3 through 5 of this report and in Reference 1. The purpose of this section is to provide the reader with a summary of what has been learned as a result of this study and therefore the justifications of the results will be limited to those required to understand the results. If the reader is unfamiliar with the ADAPT terminology, it is recommended that the first portion of Section 2 of Reference 1 be reviewed. The ADAPT presentations described there will be used to explain the results which are presented in this section.

The performance of the various classification schemes considered will be compared using the Receiver Operating Curve. This curve consists of a plot of the probability of detecting the 101 Class, P_d, versus the probability of falsely identifying a member of the 2XX Class, P_{fa}, as a member of Class 101. Figure 2.1 is an example of such a classification trade-off curve. For the Task 1 results presented on this figure, these curves are based on the assumption of Gaussian distribution for the detection statistic and the mean and standard deviation for this distribution were determined from an independent test set consisting of 85 cases. For the Task 2 studies, the experimental ROC curve based on an independent test set consisting of 1,800 cases is used. Section 4.2 of this report justifies the requirement to use the experimental ROC curve for the Task 2 data and describes the procedures for obtaining these curves. Another major difference between the comparisons of the Task 2 and Task I performance is that the Task 1 performance is evaluated using the entire 20X Class whereas the Task 2 performance is evaluated in dependently for each of the four Class 21X sub classes. This difference will tend to over-rate the performance of the Task 1 results by approximately a factor of five.

In general, the effect of the various classification laws were significantly different on each of the 21X targets. This is illustrated in the classification trade-

FIGURE 2.1 - CLASSIFICATION TRADE-OFF CURVE COMPARING THE FISHER CLASSIFICATION PERFORMANCE FOR VARIATIONS IN TARGETS, NUMBER OF PULSES AND COHERENT VERSUS INCOHERENT RADAR SYSTEMS



off curve shown in Figure 2.1. The solid lines represent the performance of the Fisher discriminant on the 211 and 214 targets. We see that the performance on the 211 target is approximately an order of magnitude better than that on the 214 target. The results presented in Section 4 of this report show that in general this is the case. Specifically, the 214 and 212 targets are always considerably more difficult to classify then the 211 and 213 targets. Although there are sometimes significant differences between all four targets for many classifiers, the 212 and 214 targets are classified with approximately the same performance. The majority of the performance comparisons presented in this section of the report will only consider the performance on the 214 target. For the detailed performance on the other three targets, the reader is referred to Sections 4 and 5.

2.1 Selection of Radar Signatures

The comparison of the performance of various classification schemes using different radar signatures has led to the recommendation that for the set of procedures examined in this study the signature used for the classification of the 101 versus 2XX targets should have the following four characteristics: 1) the signature should be from 10 or less pulses, 2) the signature should consist of the separate squares of the real and imaginary portions of the coherent return, 3) the data should not be transformed to the frequency domain, and 4) only the principal polarization of the return need be used.

The conclusion regarding the use of only 10 pulses is based on the results of the Task 1 studies. The major arguments leading to these conclusions consist of the performance of the 30 pulse versus the 10 pulse classification laws shown in Figure 2.1. In Reference 1, it is shown that this performance difference is essentially that which would be expected by the improvement in the knowledge of the statistics due to the use of three times as many cases. Thus, the classification mechanism is the same for 30 and 10 pulses. Analysis of the relative importance vectors for the 30 and 10 pulse algorithms shown in Figures 2.2 and 2.3 substantiate this conclusion. These relative importance vectors show that although a finer structure is apparent on the 30 pulse relative importance vector, this fine structure has a very noise-like character, a much smaller amplitude and is super-

FIGURE 2.2 - RELATIVE IMPORTANCE OF SIGNAL ELEMENT
CORRESPONDING TO INDEXING VARIABLE
FOR 30 PULSE INCOHERENT
CLASSIFICATION ALGORITHM

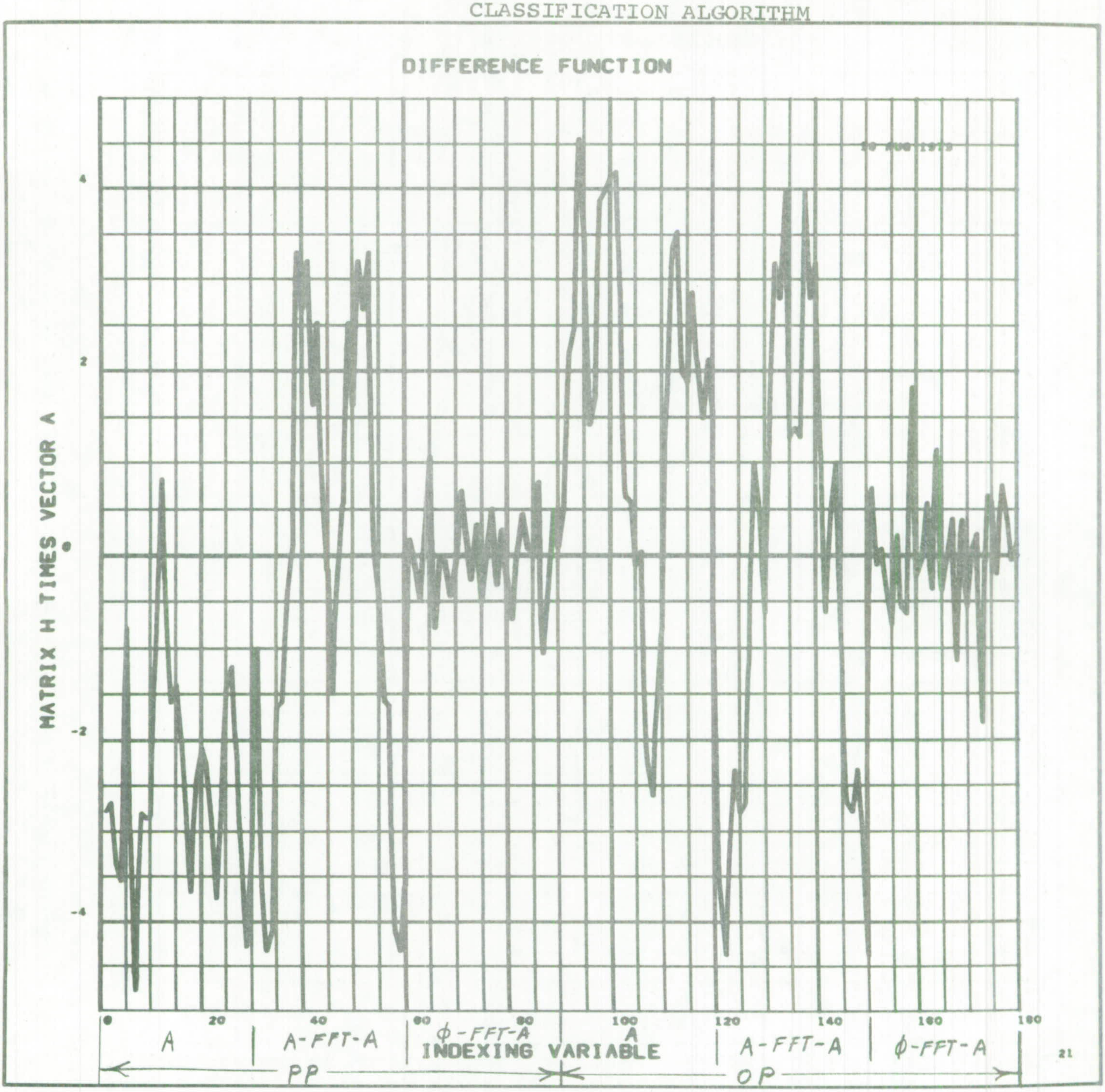
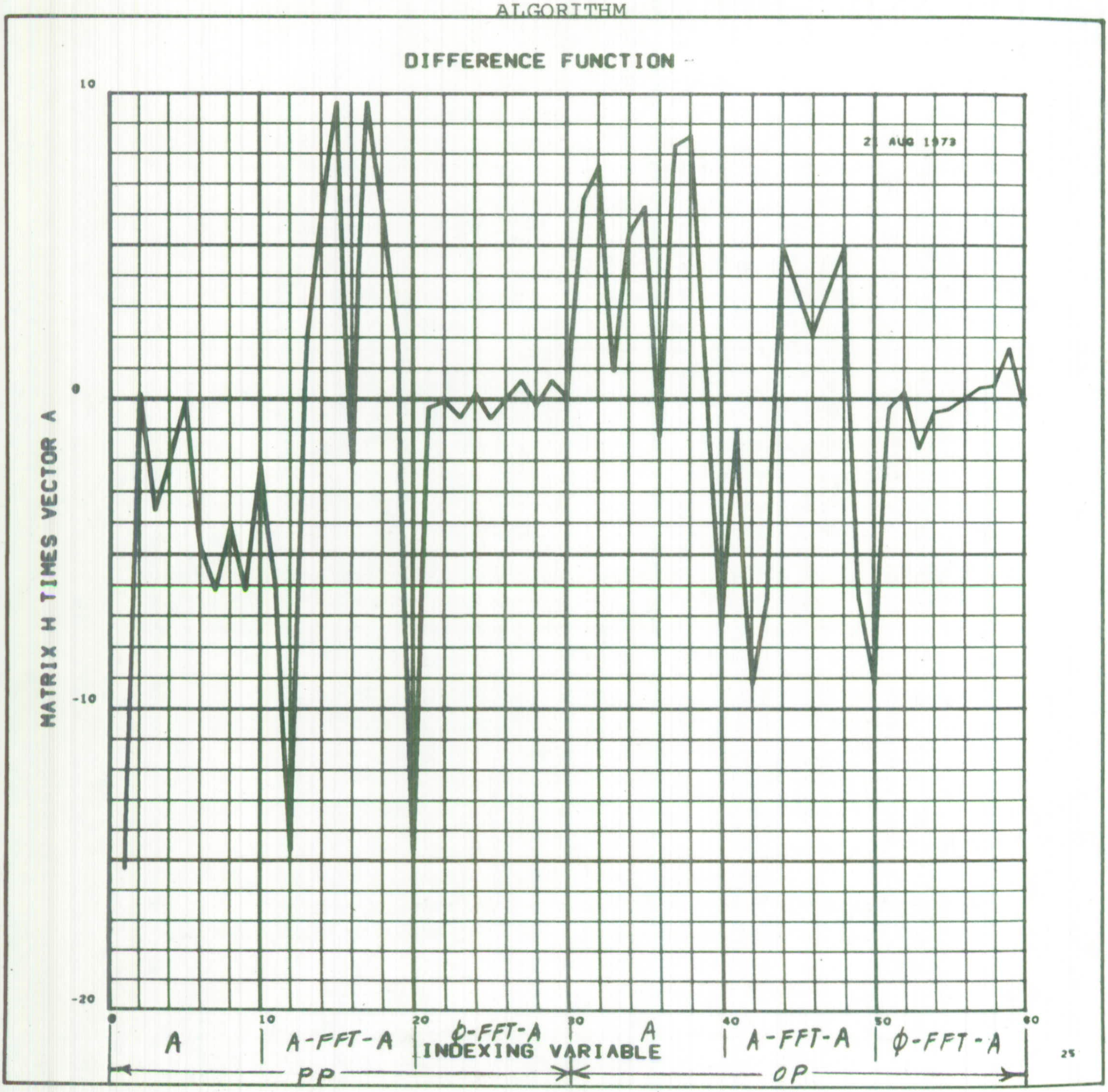


FIGURE 2.3 - RELATIVE IMPORTANCE OF SIGNAL ELEMENT CORRESPONDING TO INDEXING VARIABLE FOR 10 PULSE INCOHERENT CLASSIFICATION ALGORITHM



imposed on the same over-all structure as observed for the 10 pulse relative importance vector. Similar comparisons with the 200 pulse relative importance vector led to the conclusion that even the 200 pulse data vector did not introduce a new mechanism for the classification. Thus, 10 pulses should be adequate to obtain the physical information which is available for classification.

It is recommended that the 214 data histories be processed in an incoherent manner to provide a comparison of the coherent and incoherent processing. However, in lieu of this direct comparison, an analysis of the relative importance vector for the Fisher classification law presented in Figure 2.1 has been carried out in Section 4.5 which suggests that use of the coherent portion of the return would result in both a significant gain of total information and an increase in the performance of the classification algorithm. Thus, it is recommended that coherent processing be used.

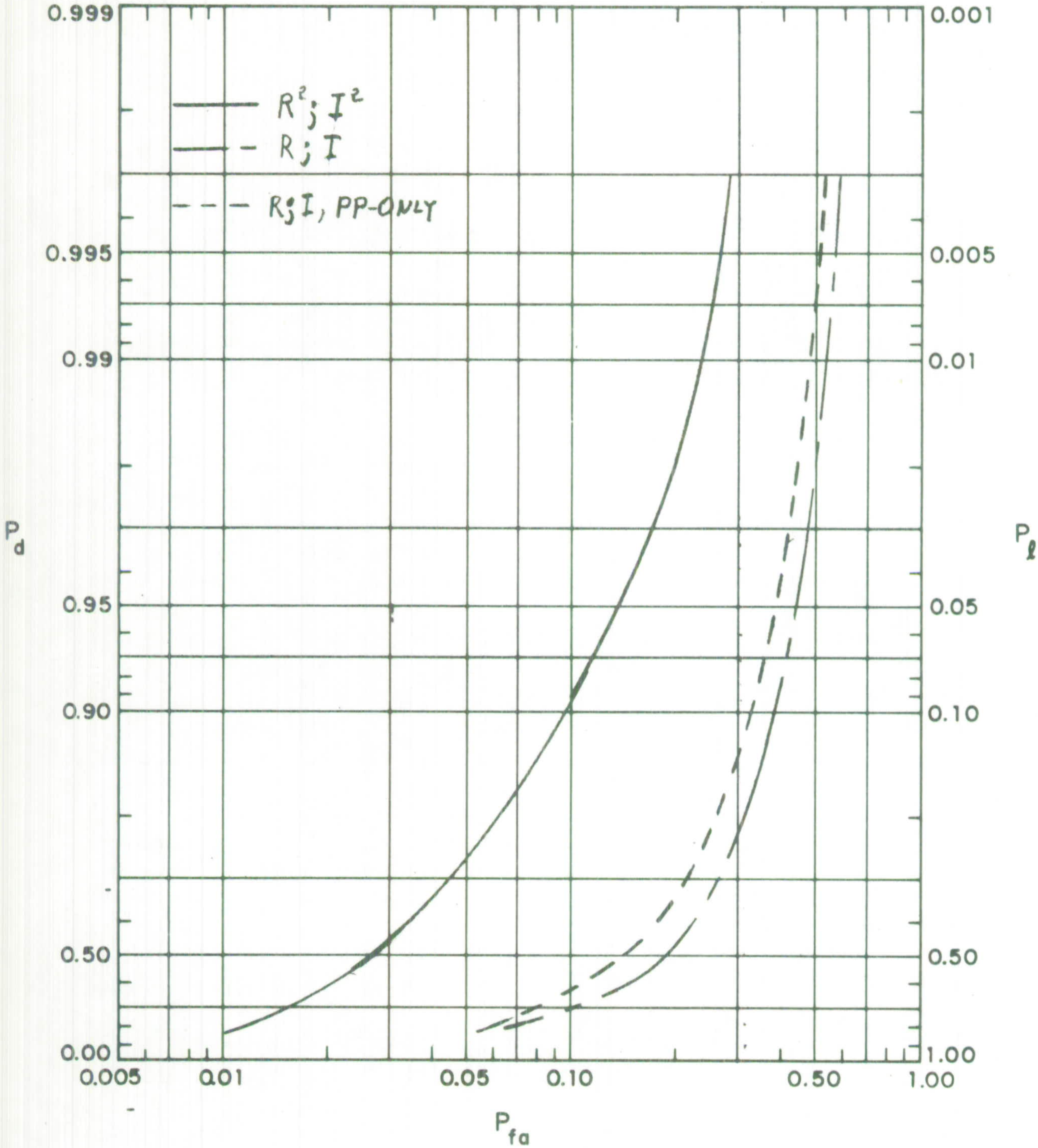
Figure 2.4 compares the performance of the minimum variation ratio classifier[1] on signatures composted of 1) the square of the real and imaginary portions of the signal and 2) the unsquared real and imaginary portion of the signal. This comparison shows an advantage for the use of the square of the real and imaginary portions of the signal. The analysis presented in Section 4.5 shows this conclusion to be valid for all of the linear classifiers. Since this requires very little additional complexity in the classification scheme and yields significant improvements in classification performance, it is recommended that the signature used for this study be composed of the separate squares of the real and imagainary components.

The second comparison shown in Figure 2.4 is between the classification performance obtained using the real and imaginary components of both the principal and orthogonal polarizations and the performance obtained using these components from only the principal polarization. This comparison shows that when the real and imaginary components of the signature are used the deletion of the opposite polarization does not significantly degrade the classification performance. Examination of the relative importance

---

1) See Section 4.1 for definition of this classification scheme which is also known in the literature as, "Simultaneous Diagonalization"

FIGURE 2.4 - CLASSIFICATION TRADE-OFF CURVE COMPARING MINIMUM VARIATION RATIO CLASSIFIER PERFORMANCE FOR SEPARATING THE 214 TARGET USING DIFFERENT PROCESSING AND PORTIONS OF THE RADAR SIGNATURES



vector for the minimum variation ratio classifier showed that this was to be expected because for the orthogonal polarization the relative importance vector had an appearance similar to the relative importance vector for the principal polarization. The same is true when one examines the relative importance vector for using the square of the real and imaginary components associated with the solid curve in Figure 2.4. This relative importance vector is shown in Figure 2.5 and one may verify the similar behavior of the principal and orthogonal polarization portions. This suggests that the results shown on Figure 2.4 for the unsquared real and imaginary components are general and that little loss in performance will occur if one restricts the signature to the principal polarization of the return.

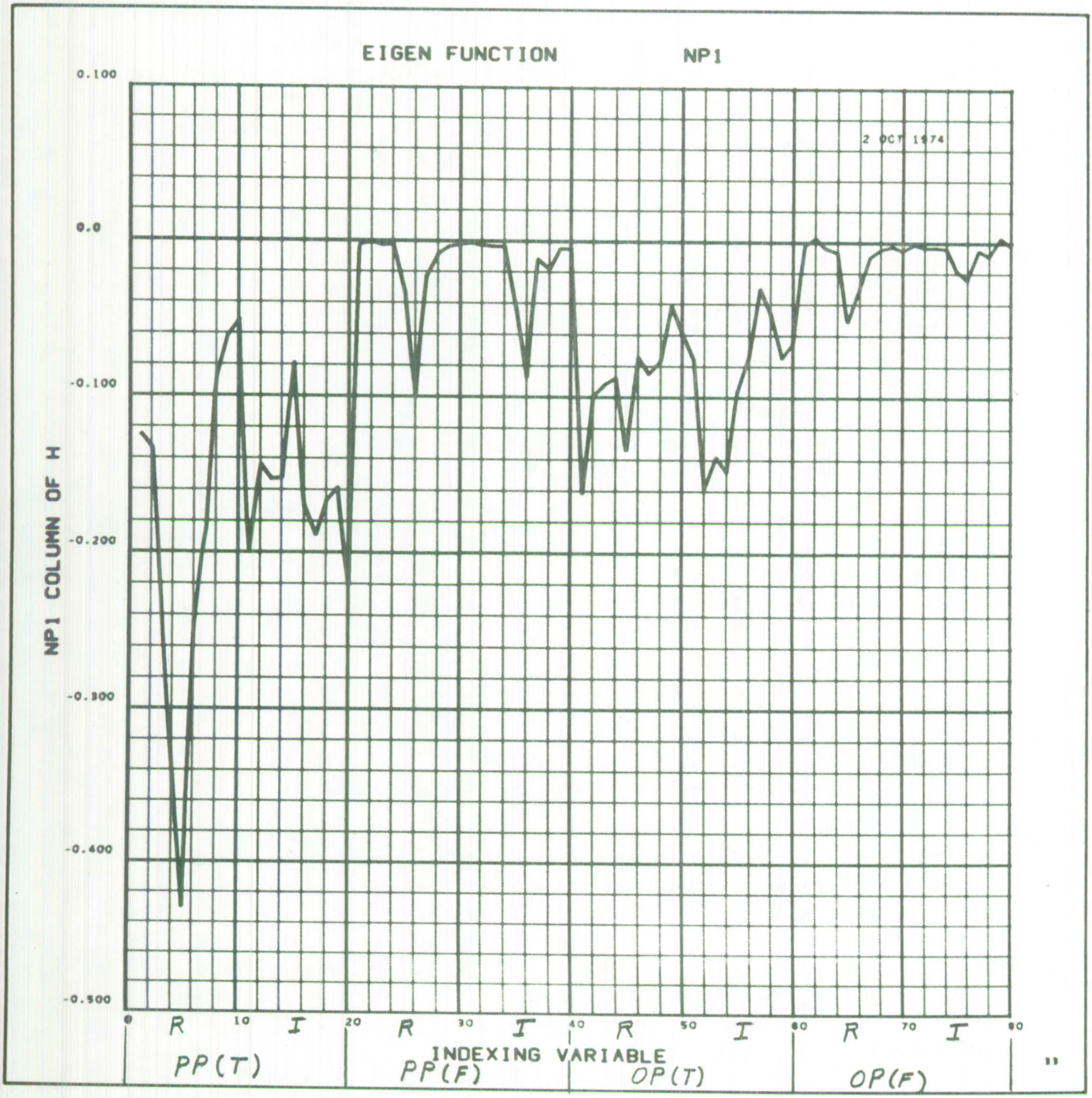
Figure 2.5 also explains why the transform to the frequency domain improved the Task 1 performance but not the Task 2 performance. When one uses the amplitude and phase as was done in the Task 1 studies, the transformation to the amplitude and phase in the frequency domain is non-linear. However, when one uses the real and imaginary components the transformation is linear. The relative importance vector, shown in Figure 2.5, is the relative importance vector based on using the signature composed of the square of the real and imaginary components. It shows that the real and imaginary components have significantly different behavior for the time domain portion of this relative importance vector but that in the frequency domain the real and imaginary components behave similarly. Thus, if one were to sum these components and utilize this sum which in the time domain is just the square of the incoherent signature used in Task 1 significant information would be lost from the time domain.

2.2 Selection of Classification Procedures

The comparison of the performance of 14 classification schemes involving non-linear classification and four different linear classification schemes[1] has led to the conclusion that the best performance can be obtained using the linear classification scheme consisting of double thresholding the projection of the data on the ADAPT optimal coordinates. By repeated application of this classification scheme, using different ADAPT optimal coordinates, it should be possible to significantly improve these results.

(1) In this report a classification scheme is considered linear if the detection statistic is derived from the pre-processed data vector in a linear manner even if multiple thresholds are used.

FIGURE 2.5 - RELATIVE IMPORTANCE OF SIGNAL ELEMENT CORRESPONDING TO INDEXING VARIABLE FOR CLASSIFICATION LAWS USING THE SQUARE OF THE REAL AND IMAGINARY COMPONENTS OF THE RADAR SIGNATURE



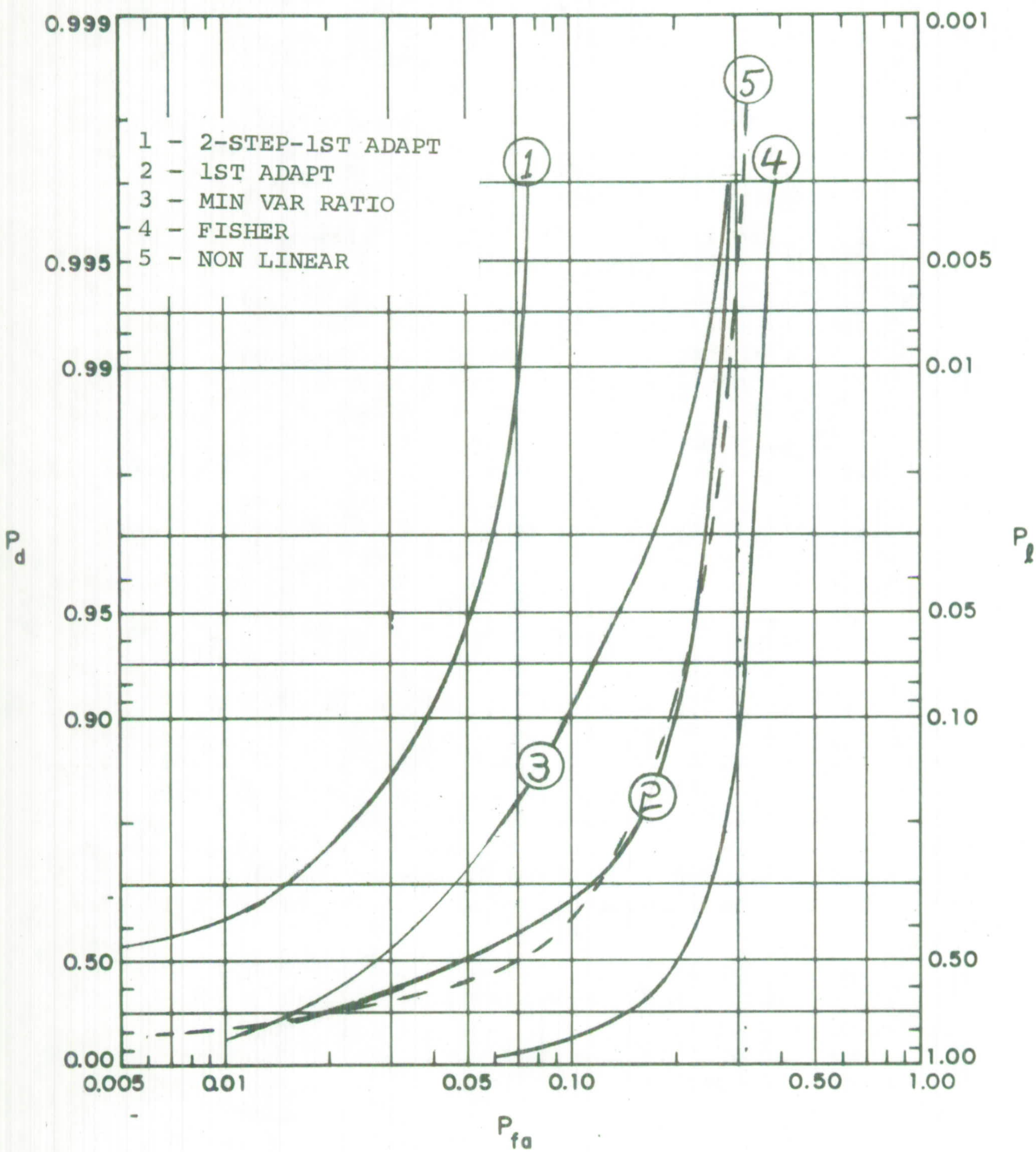
These conclusions are based on detail comparisons of large number of classification performance trade-off curves the highlights of which are summarized by the five curves which are compared in Figure 2.6. The four solid lines shown on Figure 2.6 represent the performance of four of the linear classifiers, the dash line represents the performance of the best non-linear classifier. This figure shows that several of the linear classifiers have equal or better performance then the non-linear classifier. Since the non-linear classifiers require considerable more work then the linear classifiers, it is recommended that the linear classification schemes be utilized.

The three linear classification schemes evaluated were a minimum variation ratio classifier, the Fisher classification scheme and the projection along the first ADAPT optimal coordinate direction. These comparisons show that in the region of greatest interest (i.e. $P_d \sim .99$), all three of these classifiers perform similarly. The minimum variation ratio classifier shows a slight advantage for the results presented in Figure 2.6. However, the detail analysis carried out in Section 4, shows that the projection on the first ADAPT optimal direction performs better on the other target types then do either of the other linear classification schemes. Furthermore, the projection on the first ADAPT optimal direction shows better performance as a multi-step classifier.

The two-step classification procedure consists of simply re-applying an algorithm to obtain a second estimate. If one assumes the two estimates are independent, classical sequential detection theory shows that the detection probabilities and false alarm rates obtained are simply the product of the corresponding values for each of the two decisions. The requirement for this performance estimate is that each of the steps be independent estimates of the class. In general, the two step procedure results in better performance.

The problem in using this procedure for estimating the performance is demonstrating the independence of the linear classification schemes. One natural classifier in view of the good performance of the first ADAPT optimal direction classifier is to utilize the projection on the second ADAPT optimal direction as a second classifier. The analysis of the scatter plots presented in Sections 4.1 and 4.6 suggest that the use of

FIGURE 2.6 - CLASSIFICATION TRADE-OFF CURVE COMPARING THE BEST NON-LINEAR CLASSIFIER WITH LINEAR CLASSIFIERS DERIVED USING THE RECOMMENDED RADAR SIGNATURE AND PROCESSING (FOR TARGET # 214)



the higher order ADAPT optimal directions will not significantly reduce the performance relative to the use of the first ADAPT optimal direction. In order to estimate the performance of the recommended multi-step classification procedure, it is assumed that either repeated application of the first ADAPT optimum direction classifier to an independent set of data or the use of the second ADAPT optimal direction will produce essentially the same results. These results are illustrated by the 2-Step-First ADAPT curve identified by the Numeral 1 on Figure 2.6. This curve shows significant improvement over any other linear classifiers shown. Analysis presented in Section 4.6 shows that the 2-Step procedure applied to the first ADAPT optimal direction classifier is better then the other 2-Step procedures considered.

The use of the first ADAPT optimal direction classifier in a 2-step process has the additional advantage that it may easily be extended to many more steps or a multi-step process by using even higher order ADAPT directions. The results of the scatter plot analysis presented in Section 4.1 suggests that for the present data set this can be applied to more than 15 ADAPT optimal directions. Although there will be some degradation in performance as one uses the projection on the higher order ADAPT optimal directions, this degradation should be small compared to the other uncertainties in the present analysis. Thus over the range of detection probabilities in the region of .99 the two-step classification procedure has reduced the false alarm rate to less than 10% or about 5 DB from the false alarm rate for the one step classifier.

2.3 Analysis of ADAPT Representation and Classifications

Although the major objective of this study was the definition of the best technique for separating the 101 and 2XX targets and the estimation of the performance that could be expected for this classification, the ADAPT programs also provide a mechanism for analyzing the data. As a result of this analysis much has been learned and the most important of this is summarized below:

Characteristics Observed From ADAPT Representation

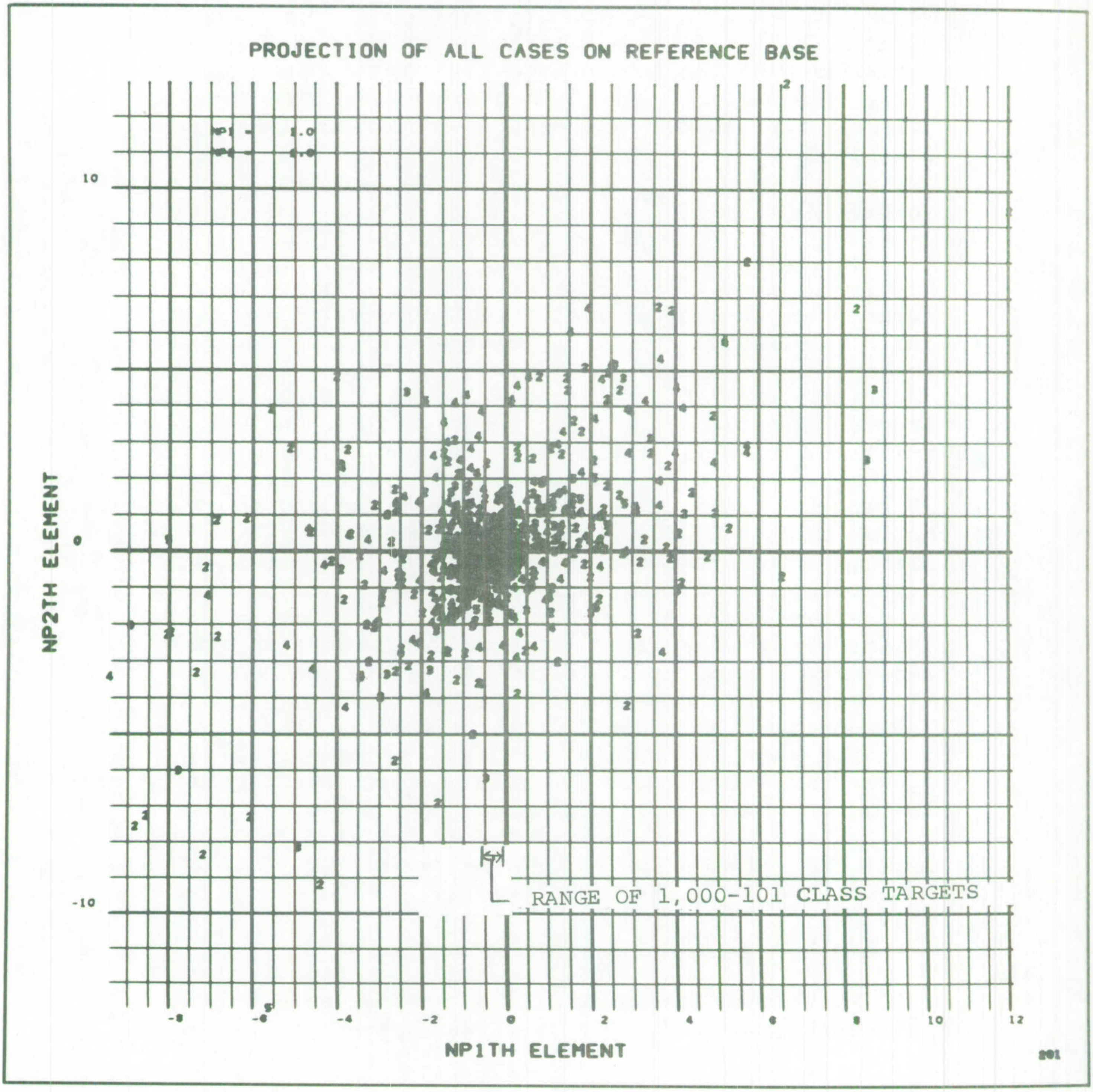
There are two results which have been observed but have not yet been explained. The first of these is that when one examines the first optimal functions and most of the relative importance vectors developed in this study one sees characteristics similar to the relative importance vector presented in Figure 2.5. Specifically, we see that the weighing of the relative importance vector in the time domain is significantly greater than that in the frequency domain. Although the relative importance vector presented in Figure 2.5 is based on using the square of the real and imaginary components, this same characteristic is true for the relative importance vectors using the real and imaginary components without the square pre-processing. This behavior is difficult to understand since one would expect the time domain and frequency domain to contain redundant information. Thus, each of these domains should be of equal importance to defining the optimal directions and to separating the 101 from the 2XX Class. These results show that this is not the case.

A second difference observed in the processing of this data is a difference between the learning data, the original test cases and the final test set. Although these two data sets should have similar statistics there were several differences in the statistical properties associated with these two data sets. The most apparent is the difference in the standard deviation of Class 101 about its mean value. The 1000-101 samples in the final test set had a standard deviation of .085 for their projection on the first ADAPT optimal coordinate while the learning and 340 case test sets each had a standard deviation of approximately .023 for their projections on its first ADAPT optimal coordinate.

Basis for Separation of Targets

The ADAPT presentation of the data also allows the explanation of the basis for the classification. Figure 2.7 shows the scatter plot projection of the 1800 independent test cases on plane defined by the first two ADAPT optimal coordinates. The Numeral 1 identifies the 101 targets and Numerals 2 through 5 identify targets 211 through 214. This plot is also typical of the plots of the 600 cases supplied for the learning and the 340 case test

FIGURE 2.7 - SCATTER PLOT PROJECTION OF 1800 INDEPENDENT TEST CASES ON PLANE DEFINED BY THE FIRST AND SECOND ADAPT DERIVED OPTIMAL COORDINATES - USING THE REFERENCE BASE



set. Examination of this plot reveals that no 101 Class targets are visible. This is because all of these targets are located at an NP1 value of -0.327 and an NP2 value of -0.18 with standard deviations of .085 and .065 in these two directions, respectively. The total extent of the 101 targets on this plot is indicated by the note and ranges from -0.01 to -0.55. The enclosing of a thousand points in this small area clearly is not within the capabilities of the plotter and thus these points appear as a single blob. Scatter plots exhibiting blow-ups of this region of space are presented in Section 3 for the learning data.

Generality of Bases

Several studies were performed to demonstrate that the base derived using this learning data should be adequate for developing classification laws to separate almost any similar targets of the 101 and 21X classes. The generality of the base for the 101 Class targets follows immediately from the small variation associated with these targets. Because of this extreme variation, the characteristics of the optimal base for representing the combined 101 and 21X targets are almost entirely determined by the 21X targets. This was verified by deriving a base without the 101 targets. Analysis of this base showed it to be essentially the same as the base derived including the 101 targets. A similar analysis was performed by re-deriving a base deleting the 214 target. This base showed slightly less similarity to the reference base than the base derived deleting the 101 Class. Thus, we conclude that the addition of other 21X targets and similar 101 signatures will not significantly effect the characteristics of the base and its ability to derive classification algorithms. Furthermore, because the recommended algorithms are based on the significantly greater variation in the to 21X Class then the 101 Class, the specific algorithms should also be applicable for a large variety of additional targets in each of these classes.

2.4 Recommendations for Additional Analysis

Additional studies would add to the usefulness of the results obtained. Six such studies are:

1) Verification of the requirement that within the processes considered in this study coherent processing should be used by re-deriving the recommended classification laws using the square pre-processing and summing the real and imaginary components to provide a signature equivalent to the square of the incoherent return.

2) Evaluate the performance of the Fisher discriminant when the mean value of Class 101 is subtracted from each target in the ADAPT optimal space and the resulting difference is squared. Note, that this analysis was carried out in the original data space but it should be significantly more effective in the ADAPT optimal space.

3) Repeat studies presented here using a smaller number of pulses to assess the possibility that satisfactory results may be obtained with even less than 10 pulses.

4) Investigate the differences between the original and final test case data sets which were supplied for this study. One approach to accomplishing this is to develop linear classification laws to separate these two sets of data and then examine the resulting relative importance vector.

5) The algorithms developed here should be evaluated in their multi-step mode against independent test cases. Realistic evaluation of this performance would require many more independent test cases then used for the present study. Thus, it is recommended that this evaluation be carried out for only one 21X target and that target should be the most difficult target expected. The procedure would be to first examine all of the available 21X targets to select the most difficult target using a relatively small number of independent test cases. Then the multi-step procedures recommended here should be applied to approximately 100,000 independent test samples for this target and the 101 target. This would provide verification of the ultimate performance which has been suggested by these studies.

6) Investigate the physical significance of the fact that the mean value of the projection of the 101 Class in the Karhunen-Loeve space differed from zero by an amount large compared to standard deviation of the 101 Class.

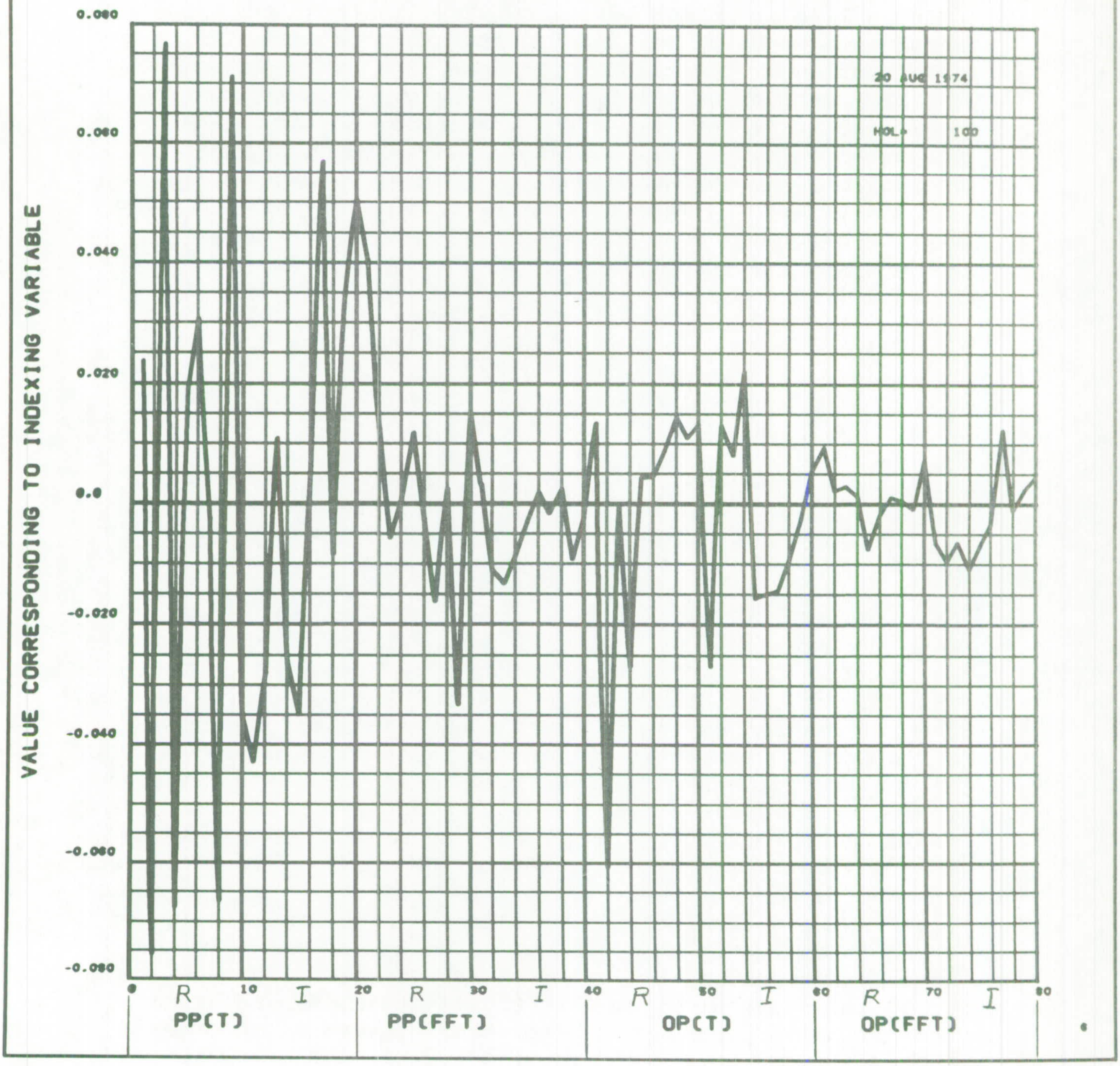
3.0 ANALYSIS OF ADAPT DERIVED OPTIMAL REPRESENTATIONS

This section of the report will present the properties of the ADAPT derived optimal bases which are used for the analysis. The information presented will be parallel to that presented in Section 3 of Reference 1 for the Task 1 bases. The results will be presented with a minimum of description since it is assumed that the reader is familiar with the information presented in Reference 1 in general and in particular in Section 3 of Reference 1.

The optimal bases used for the present study may be viewed as variations of a reference optimal base. The first base to be presented will be this reference base. This will be followed by the presentation of the variations of this base. The description of the bases will include a brief discussion of the rational and general characteristics of the base and plots showing the characteristics of the base. In general, the following information will be given for each base: 1) typical data vectors, 2) the average of all of the data vectors used to make the base, 3) the information energy or explained variation as a function of the number of terms retained in the base, 4) the optimal functions to be used in the generalized Fourier expansion, which may also be viewed as the transformation vectors to transform the data vectors to an optimal coordinate system, and 5) scatter plots of the coefficients of the dominant optimal functions. All of these bases have also been delivered to Lincoln Laboratory on tape and this section should serve as a guide to the use of the transformations on this tape.

In this study, the non-linear pre-processing did not include the equalization of the data histories as was done in Task 1. Equalization is accomplished as described in Section 3.1 of Reference 1 and results in data vectors whose values lie between 1 and 2. The major reason for introducing this equalization is to eliminate the arbitrariness which is introduced by selecting different units for different sub sets of variables. For example, if a data vector includes both amplitude and phase the relative effect of the phase would be quite different if it was measured in seconds instead of degrees. For the present study, this problem does not exist since all of the data has the same units. The other non-linear pre-processings used vary between the bases and

FIG - 3.1 SAMPLE 10 PULSE DATA HISTORY FOR TARGET 101



will be discussed with the individual bases.

3.1 Reference Base

The reference base for this study consisted of the base made up using data from ten radar pulses. The first ten components of this base were the real portion of the principal polarization of the return. The next ten components were the imaginary portion of the principal polarization of the return. Components 21 through 30 are the real portion of the Fourier transform of the principal polarization of the return and components 31 through 40 are the imaginary portion of this Fourier transform of the principal polarization of the return. Components 41 through 80 are the same components except they are obtained from the opposite polarization. Typical data histories for the two classes of targets considered in this study are presented in Figures 3.1 and 3.2. One hundred samples representing different initial conditions were taken from the Class 1 data and 40 samples were taken from each of the four Class 2 targets making a total set of 260 data histories to construct the reference base. The average of these 260 histories is presented in Figure 3.3.

The average of all of the data histories presented in Figure 3.3 was subtracted from each data history and then this zero mean set of data histories was processed through the ADAPT programs to derive the optimal empirical orthogonal functions (in the Karhunen-Loeve sense) for representing this ensemble of 260 data histories. The amount of variation which is explained by each term in the optimal representation is shown in the information energy plot presented in Figure 3.4. This figure consists of two curves. The lower curve presents the information energy in each of the terms and the upper curve is the cumulative sum of the information energy in all of the preceding terms. Thus, the first point on this curve indicates that the first term in the optimal representation explains approximately 20% of the information contained in the data set. The second point on the lower curve occurs at a value of approximately 12%. This indicates that the second term in the optimal expansion contains approximately 12% of the information in the data set. The second point in the upper curve occurs at a value of approximately 32% (the sum of 12 and 20) indicating that the first and second terms of the optimal expansion taken together explain 32% of the

FIG - 3.2 SAMPLE 10 PULSE DATA HISTORY FOR TARGET 211

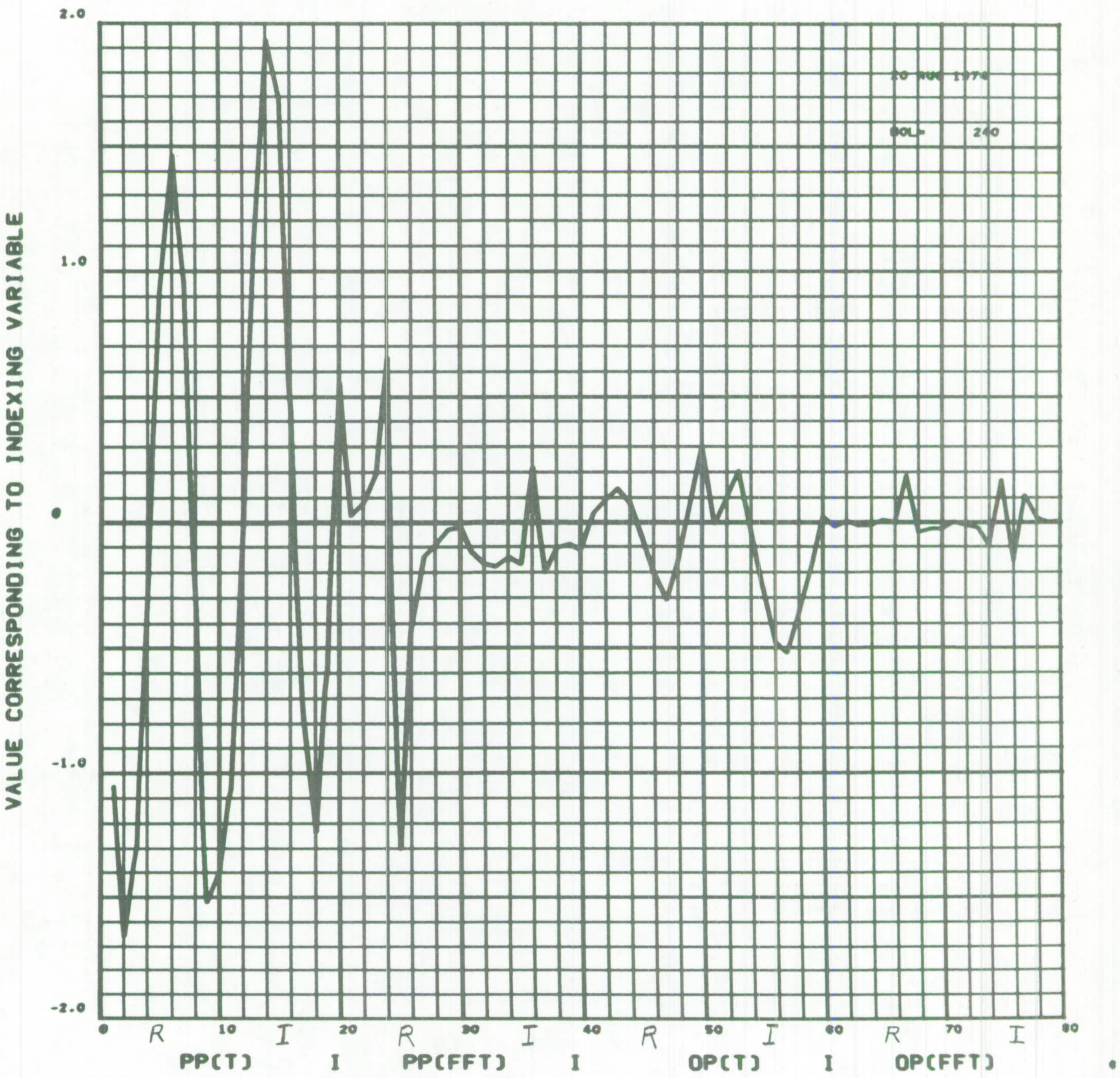


FIGURE 3.3 - AVERAGE OF ALL 260 DATA HISTORIES USED TO CONSTRUCT THE REFERENCE BASE

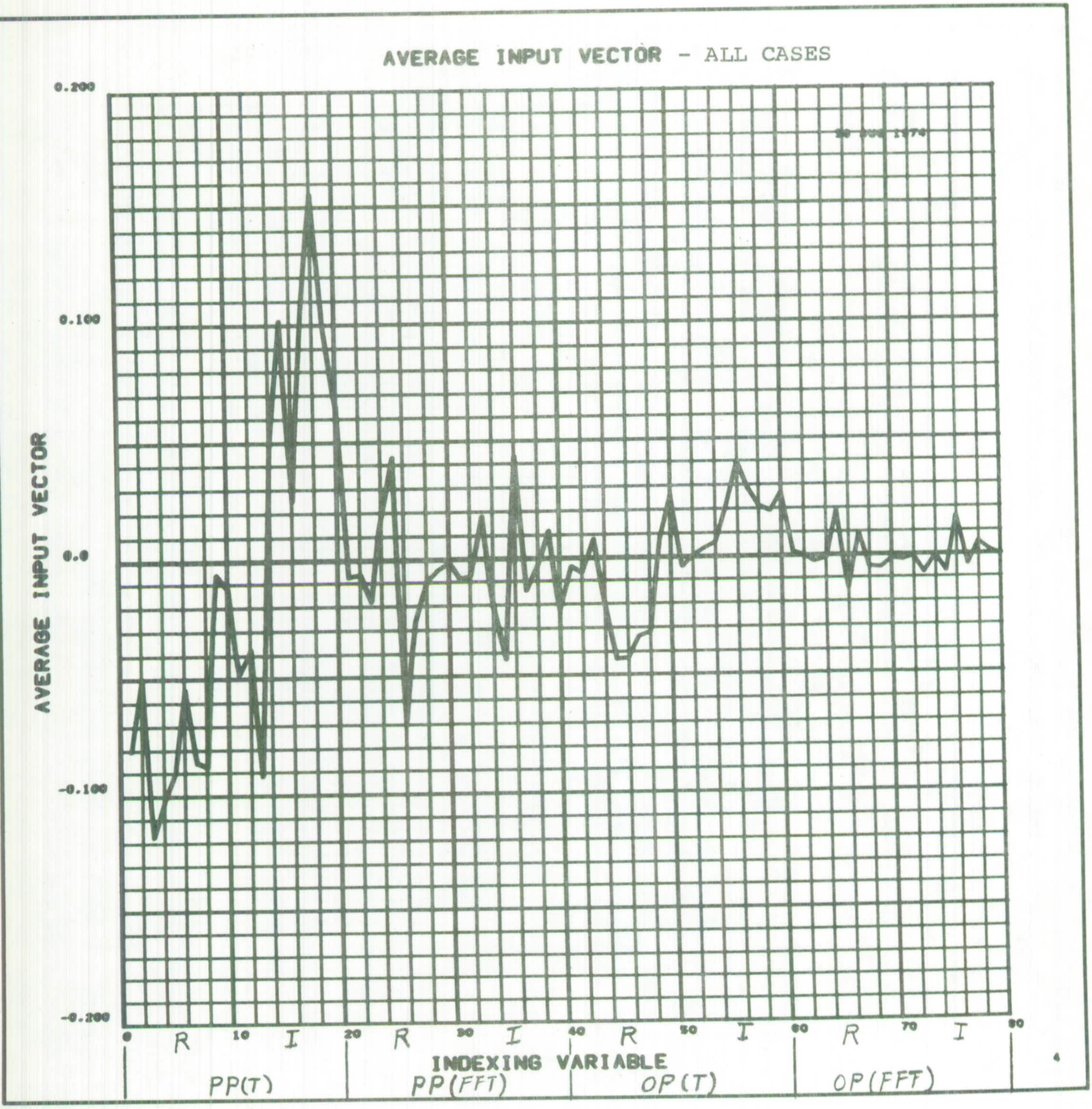
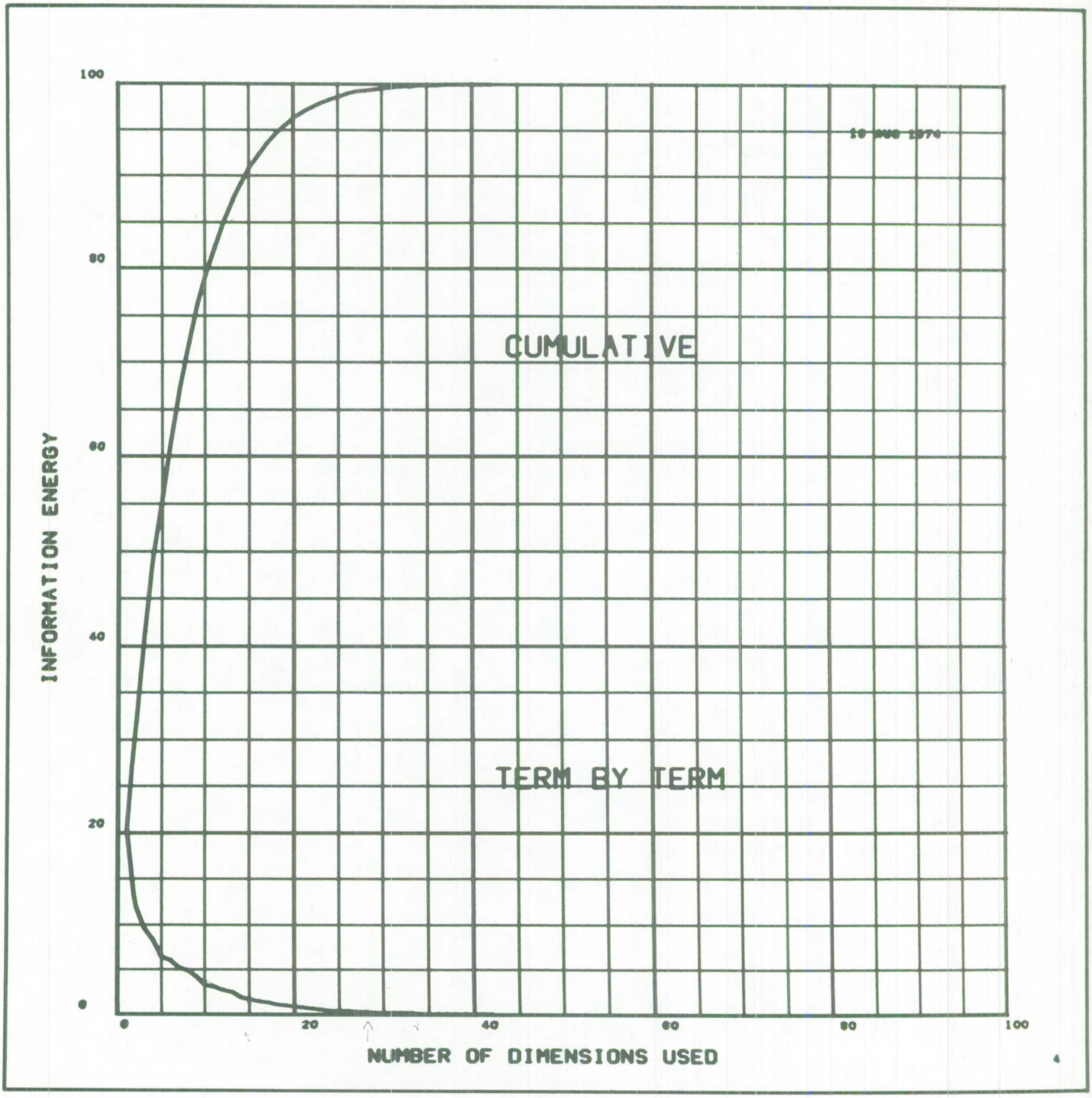


FIGURE 3.4 - EFFECT OF DIMENSIONS ON AVAILABLE INFORMATION FOR REFERENCE BASE



information in the original set of data. Examination of the lower curve suggests that changes occur in the slope at dimensionalities of 15, 28 and 33. Thus, these dimensionalities represent candidate dimensionalities for analysis. The present study was performed at 15 dimensions which proved adequate for all of the analysis carried out.

Figures 3.5 and 3.6 present the first two optimal functions derived for this data set. These functions may also be considered as the first and second columns of the transformation matrix between the original eighty dimensional data space and the first two dimensions of the optimal space. The indexing variables shown on these optimal functions are exactly the same as the indexing variables for the typical data history shown on Figure 3.1. It is interesting to note that there is considerably more variation in the time domain variables; that is, variables 1 through 20 and 41 through 60 than in the frequency domain as defined by the remaining variables.

Figure 3.7 gives the scatter plot of the coefficients of the first versus the second term in a generalized Fourier series representation using the optimal functions presented in Figures 3.5 and 3.6. This figure may also be interpreted as the projections of the learning data on the plane defined by the first and second optimal directions of the optimal orthogonal space. On this figure the 101 targets are identified by the numeral 1 and the four 21X targets are identified by numerals 2 through 5. The reader will quickly note that no numeral ones are visible. This is because the numeral ones are buried in a single point located near the center of this scatter plot. This is characteristic of all the scatter plots in this study. Physically, it is a manifestation of the characteristic of the 101 target that it has considerably less variation then any of the 21X targets. The discrimination analysis presented in other sections will show that this is a very effective discriminant for separating these two classes of targets.

The remainder of the transformation from the original data space to the 15 dimensional optimal space used for analysis in this study are presented in Figures 3.8 through 3.20. The reader should note that although the difference is less significant than in the first two optimal functions, in general one still finds considerably more variation occurring

FIGURE 3.5 - FIRST OPTIMAL FUNCTION REFERENCE BASE

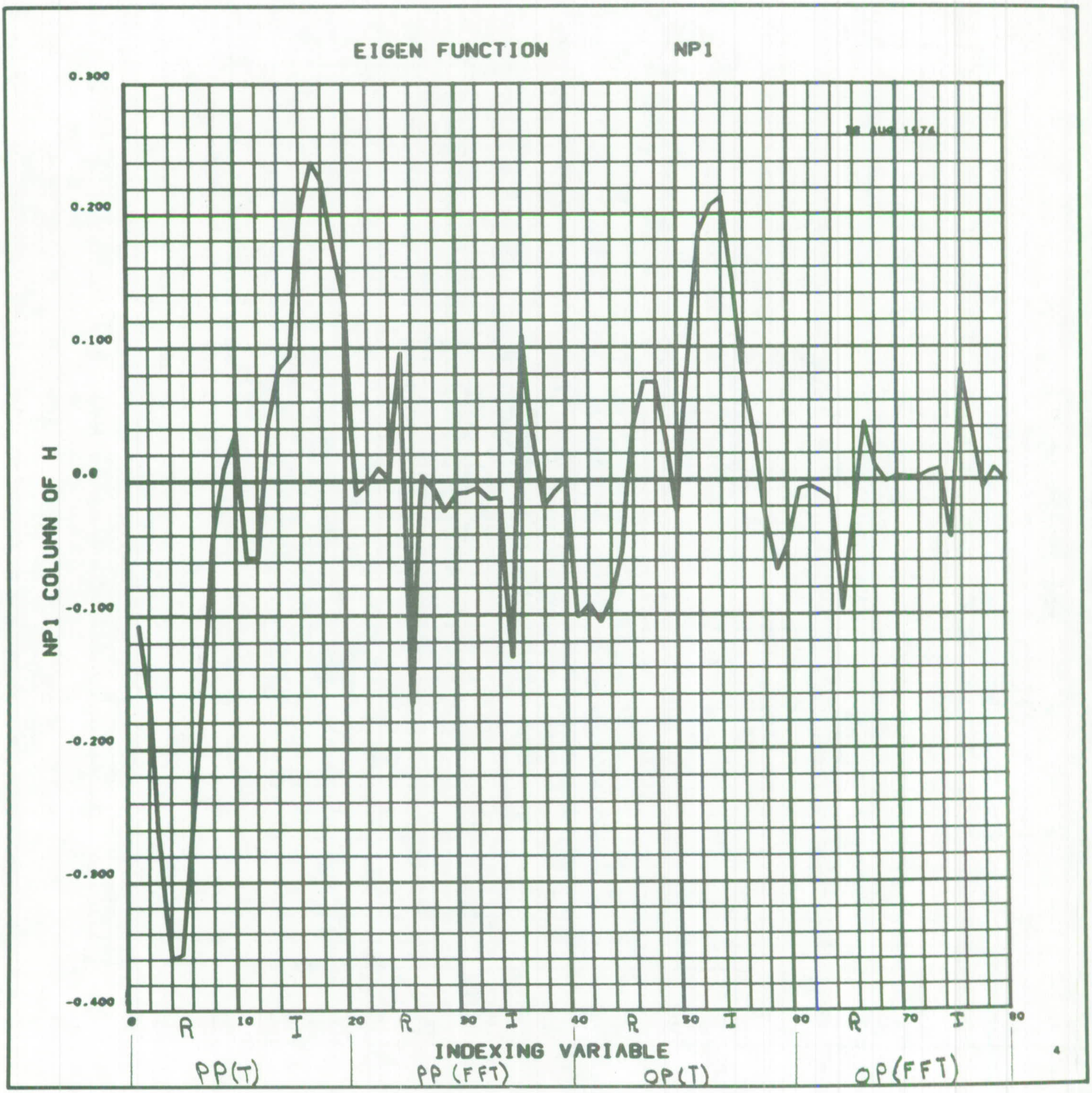


FIGURE 3.6 - SECOND OPTIMAL FUNCTION REFERENCE BASE

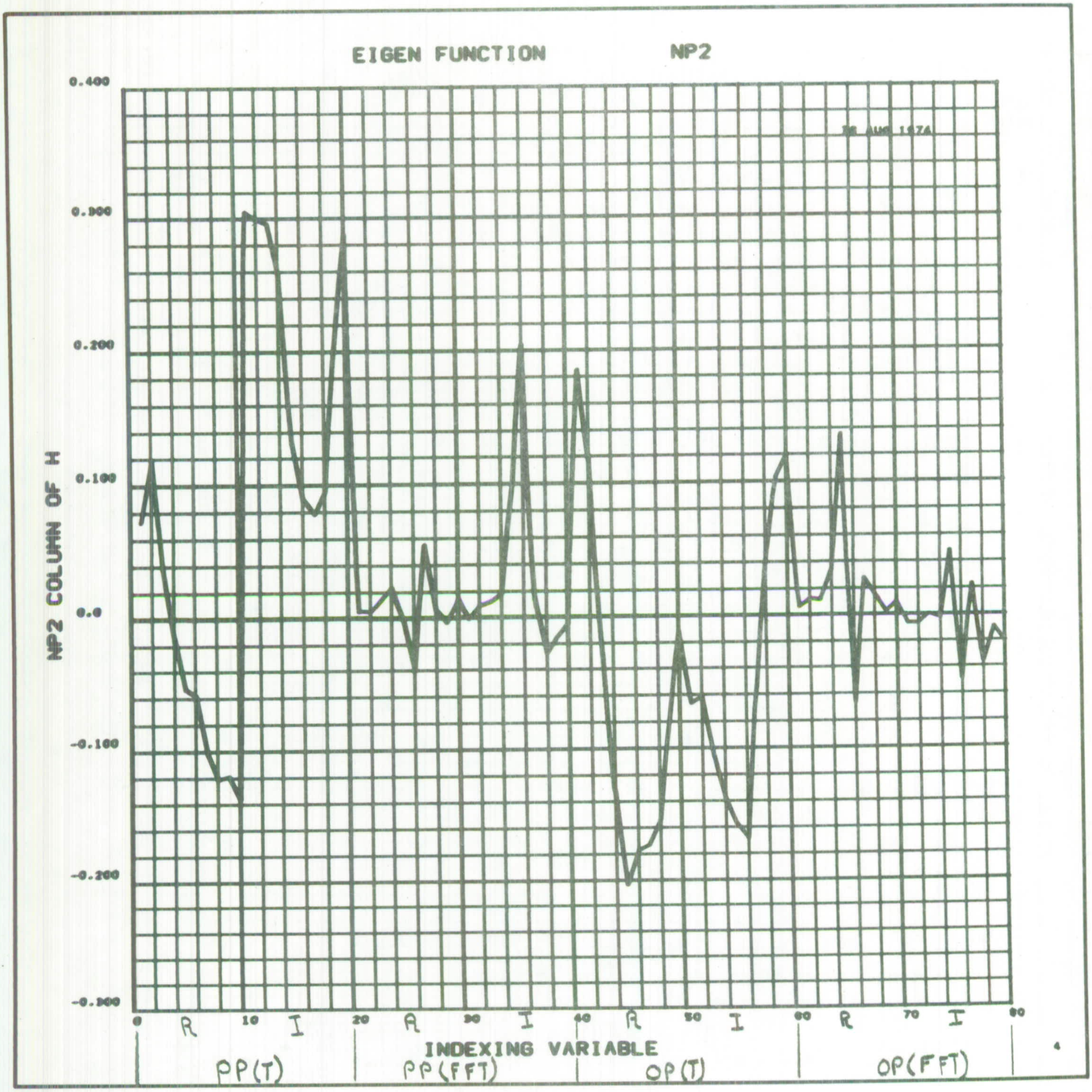


FIGURE 3.7 - SCATTER PLOT OF COEFFICIENTS OF FIRST VERSUS SECOND TERMS IN OPTIMAL REPRESENTATION - REFERENCE BASE

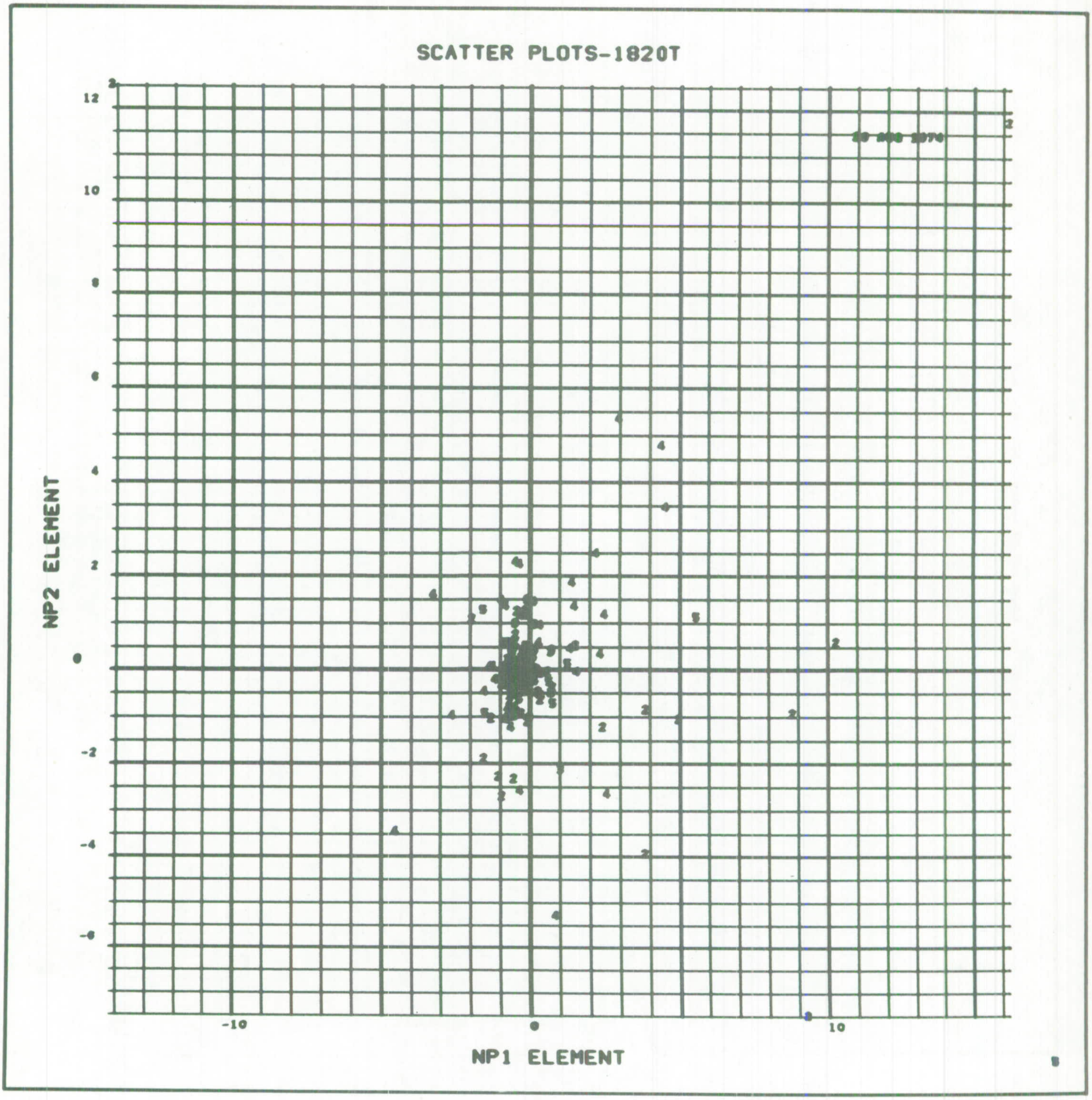


FIGURE 3.8 - THIRD OPTIMAL FUNCTION REFERENCE BASE

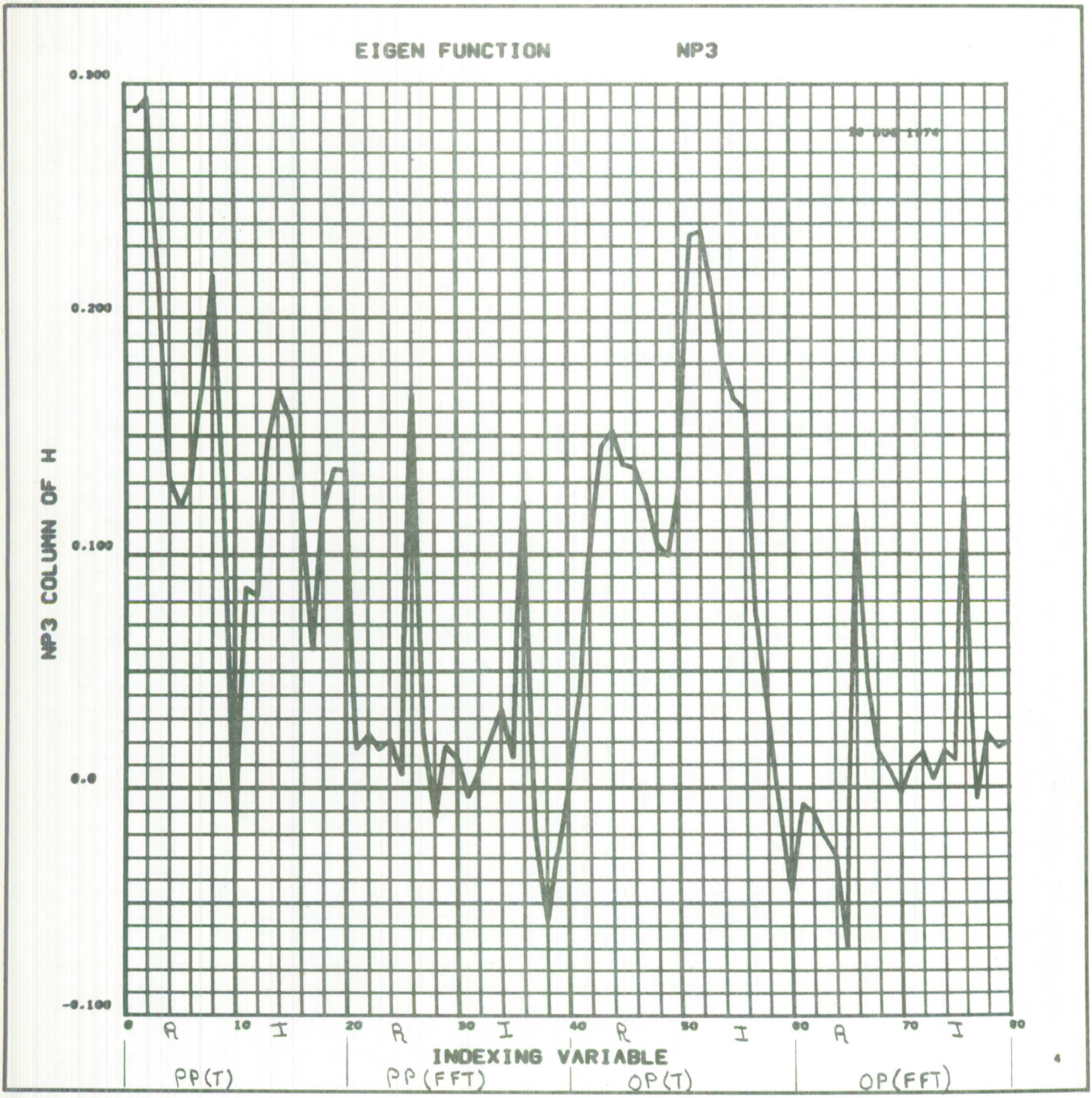


FIGURE 3.9 - FOURTH OPTIMAL FUNCTION REFERENCE BASE

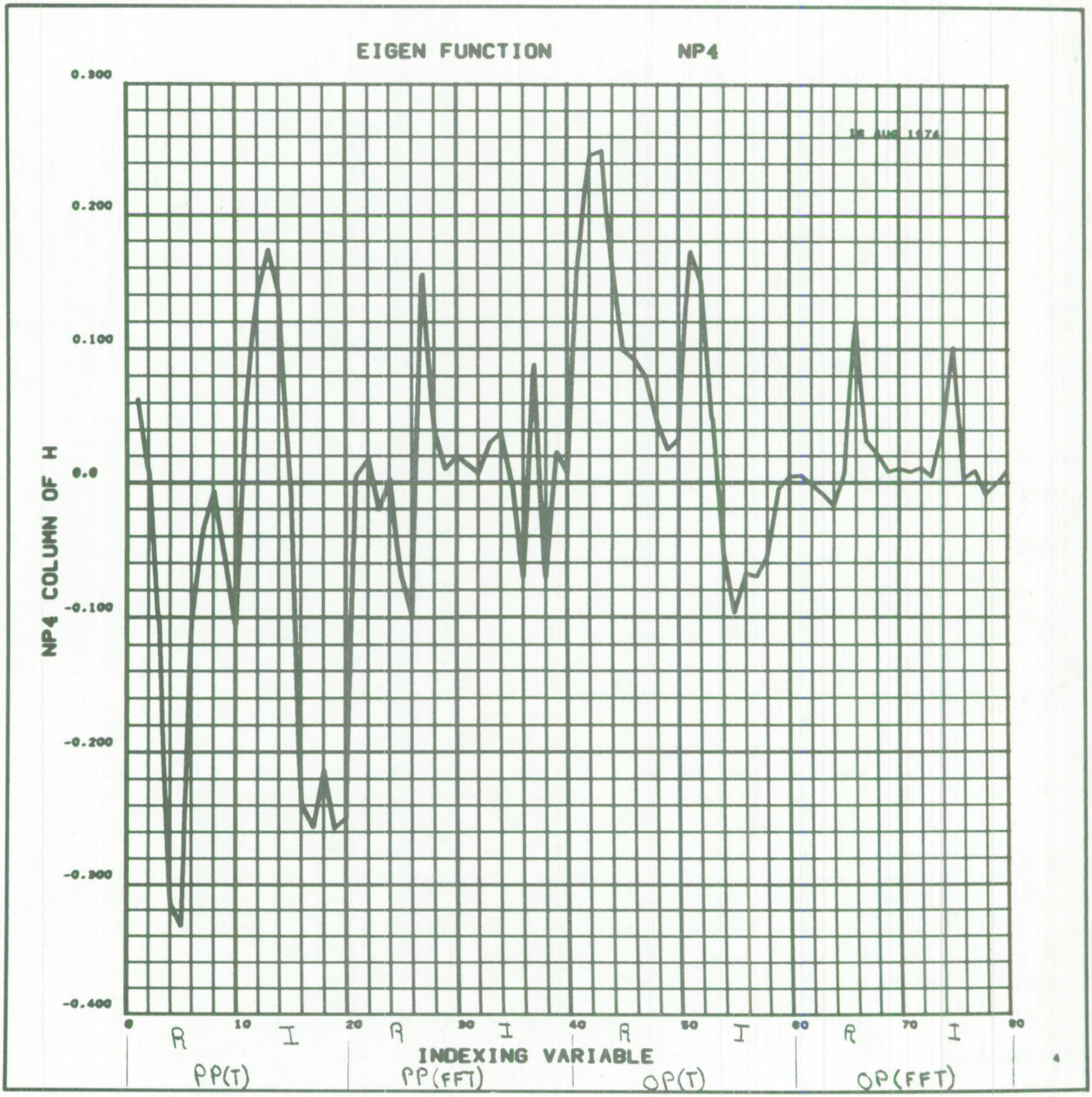


FIGURE 3.10 - FIFTH OPTIMAL FUNCTION REFERENCE BASE

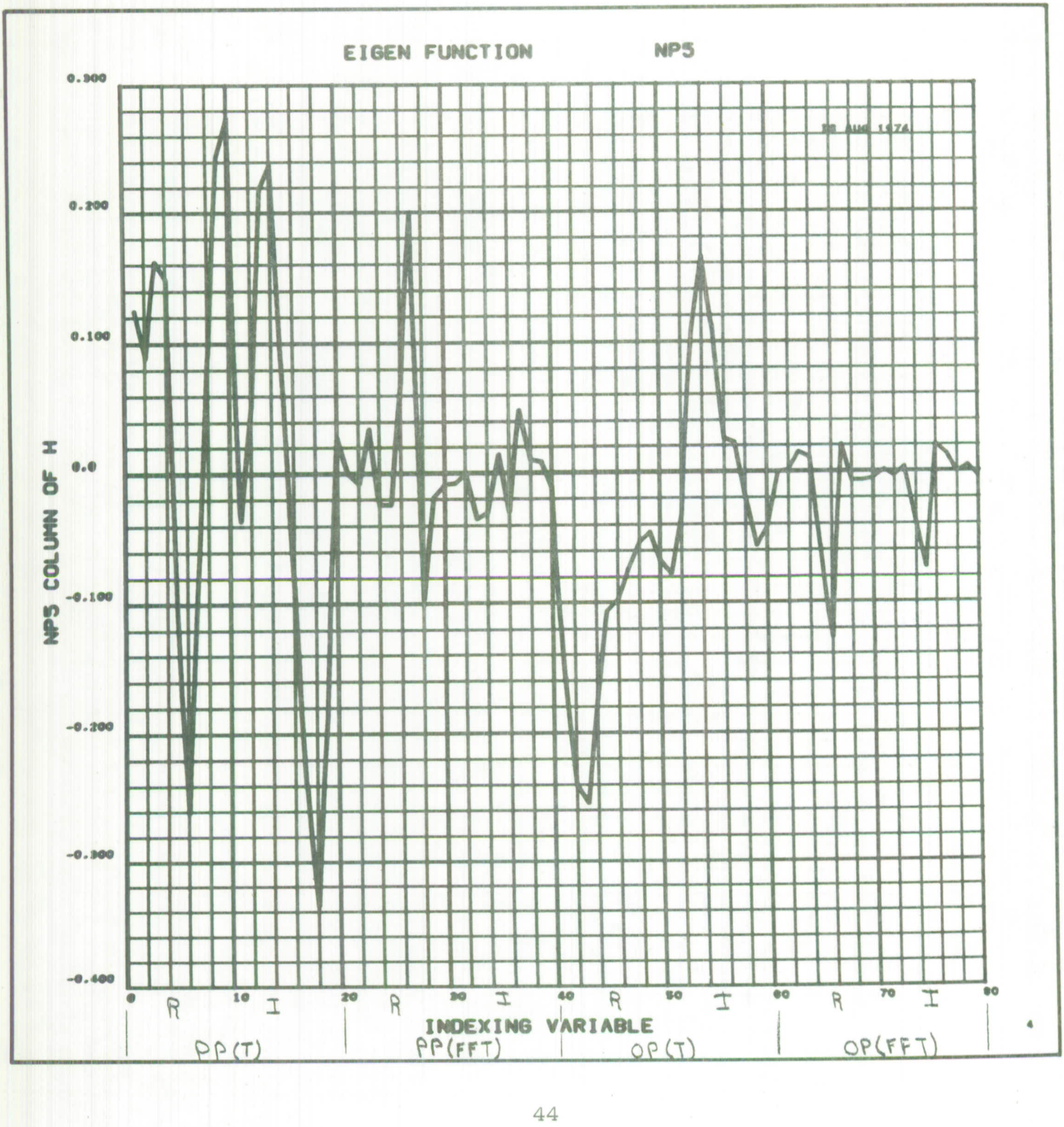


FIGURE 3.11 - SIXTH OPTIMAL FUNCTION REFERENCE BASE

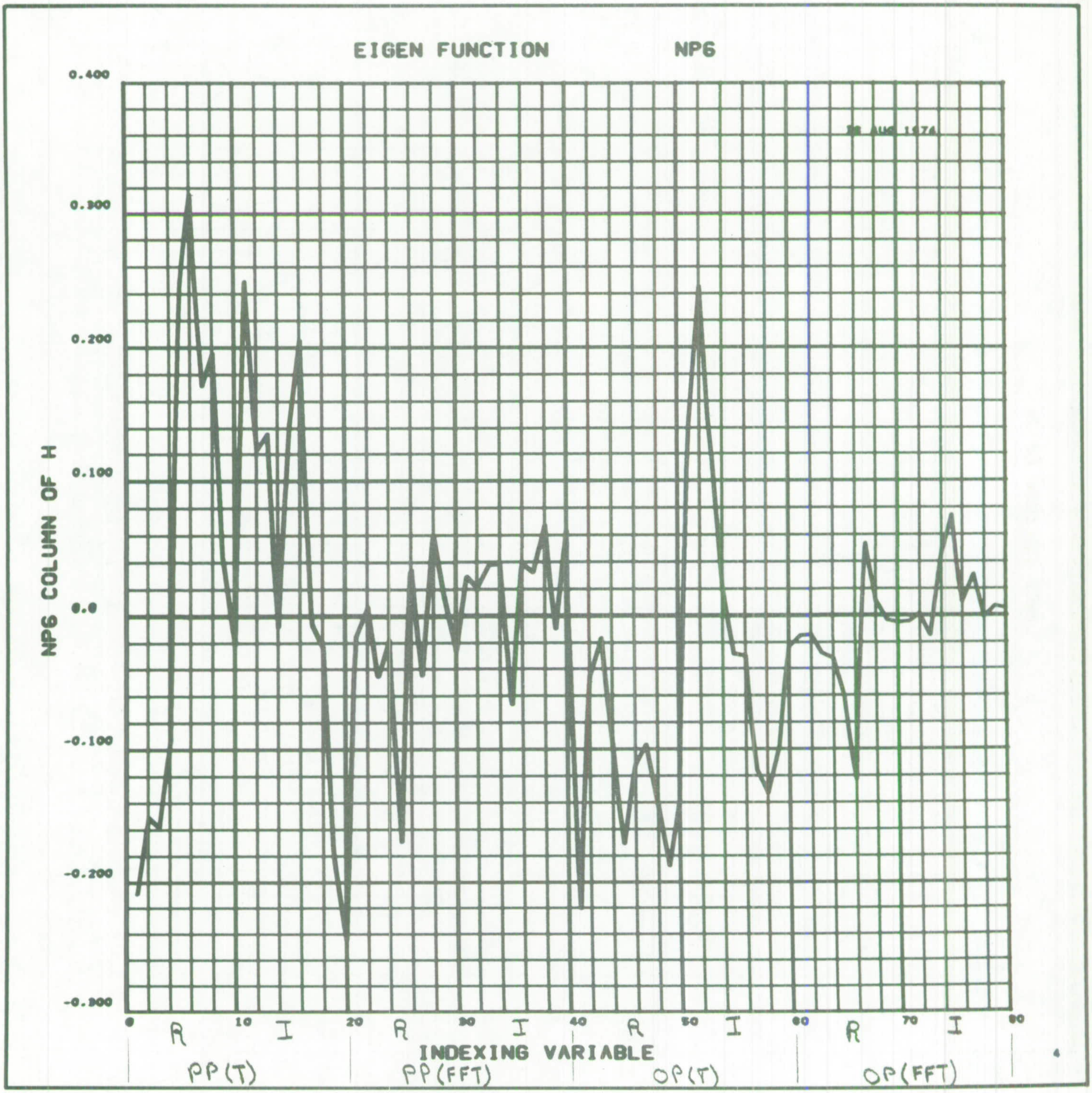


FIGURE 3.12 - SEVENTH OPTIMAL FUNCTION REFERENCE BASE

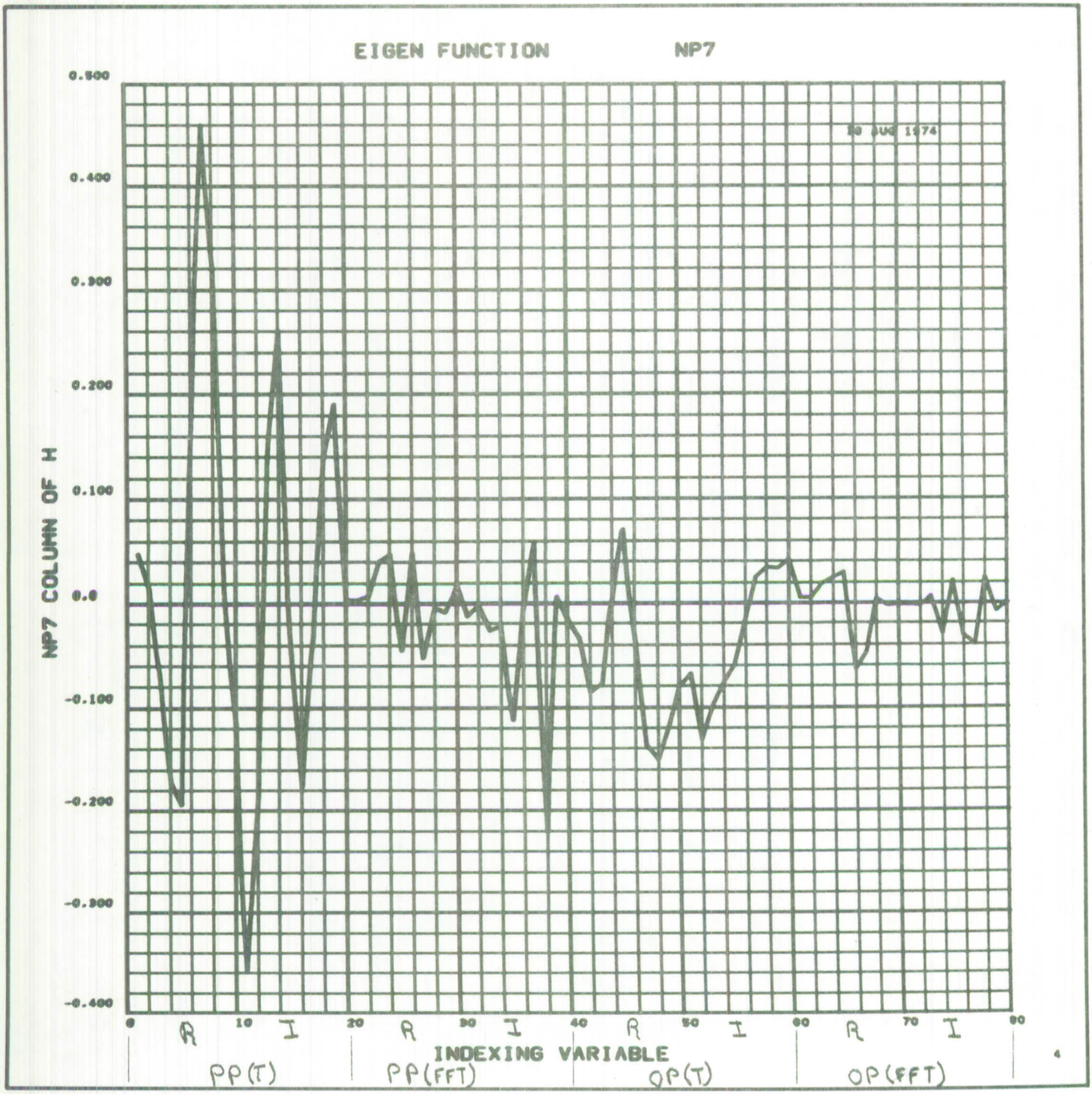


FIGURE 3.13 - EIGHTH OPTIMAL FUNCTION REFERENCE BASE

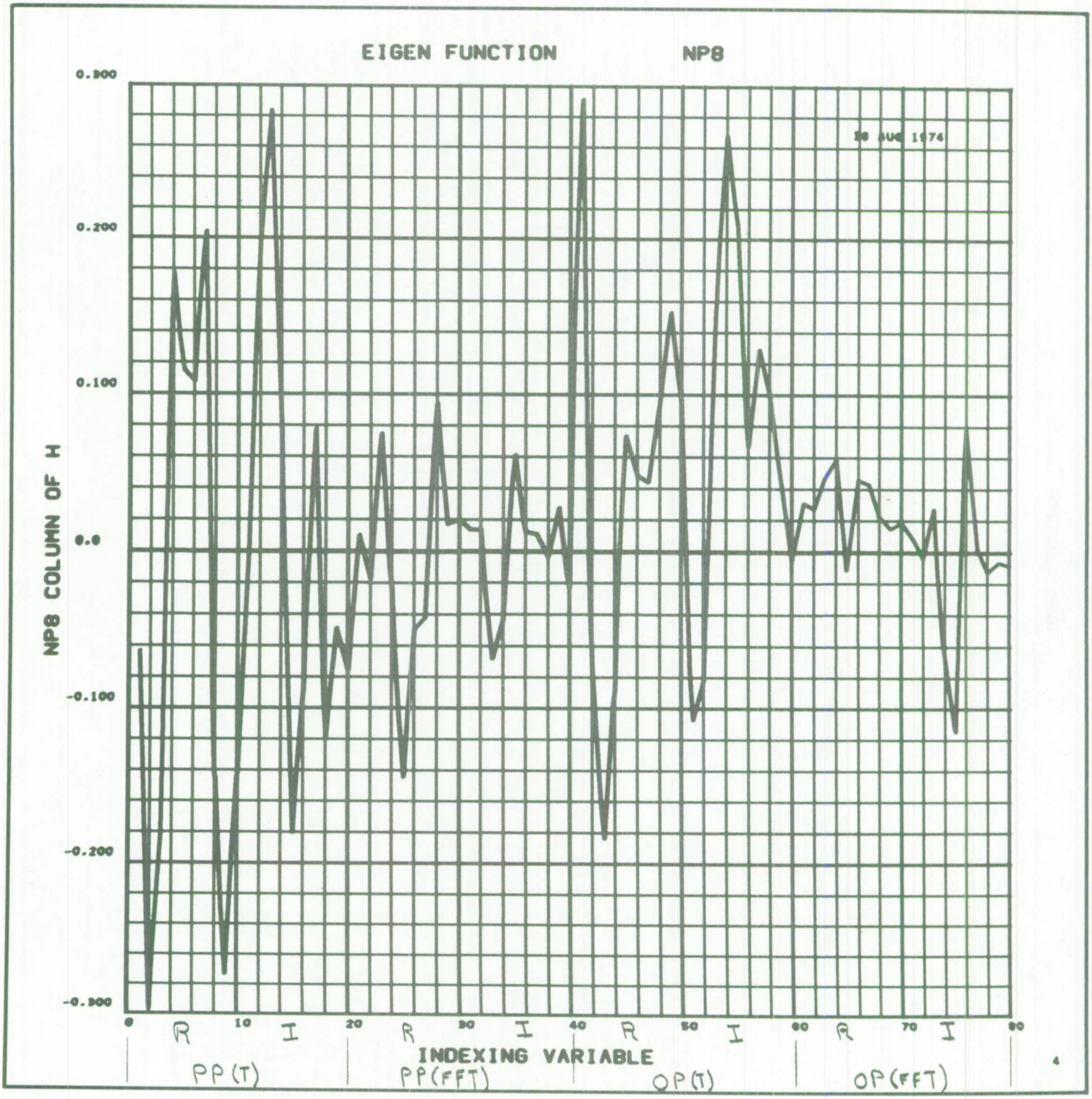


FIGURE 3.14 - NINTH OPTIMAL FUNCTION REFERENCE BASE

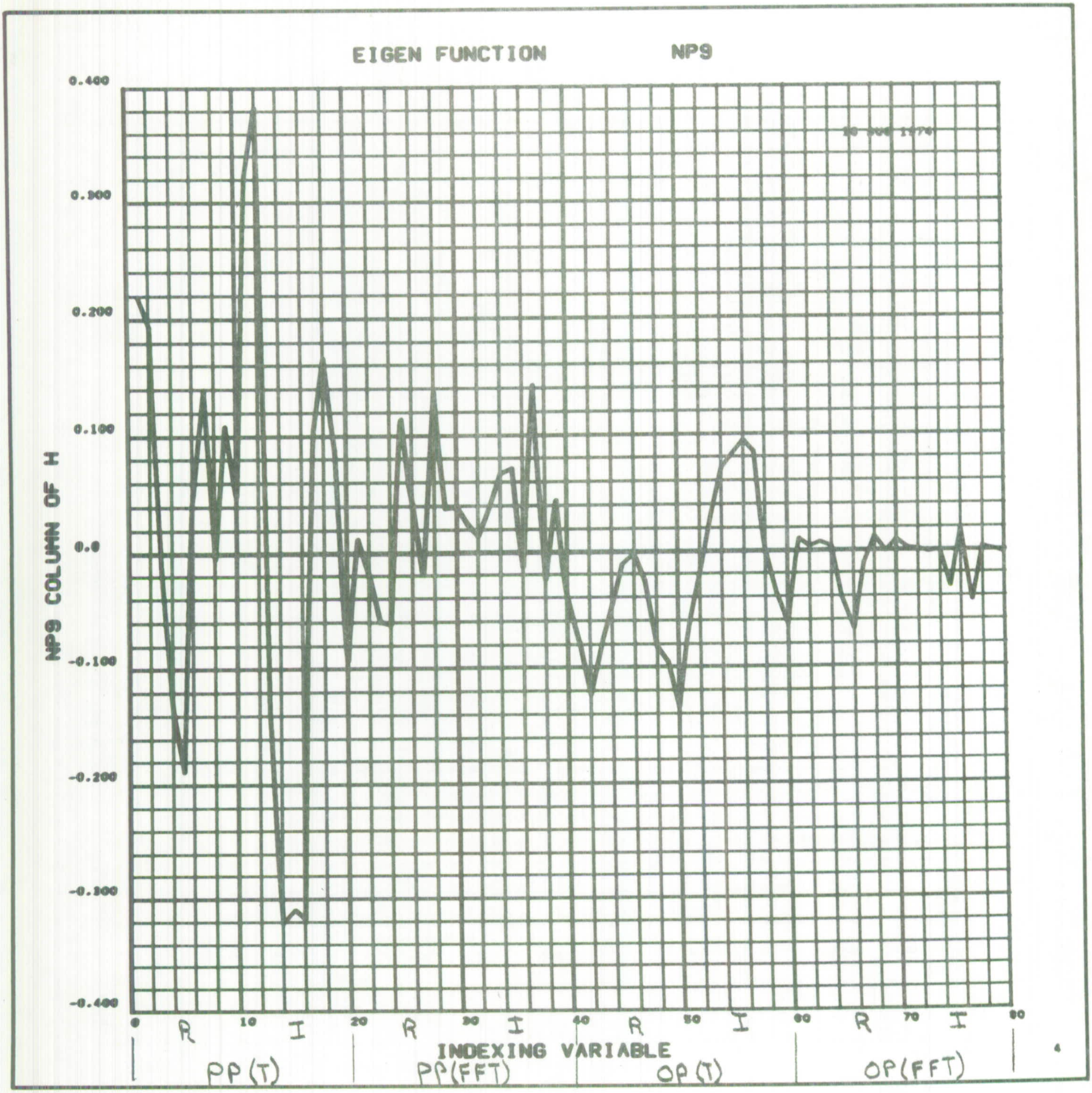


FIGURE 3.15 - TENTH OPTIMAL FUNCTION REFERENCE BASE

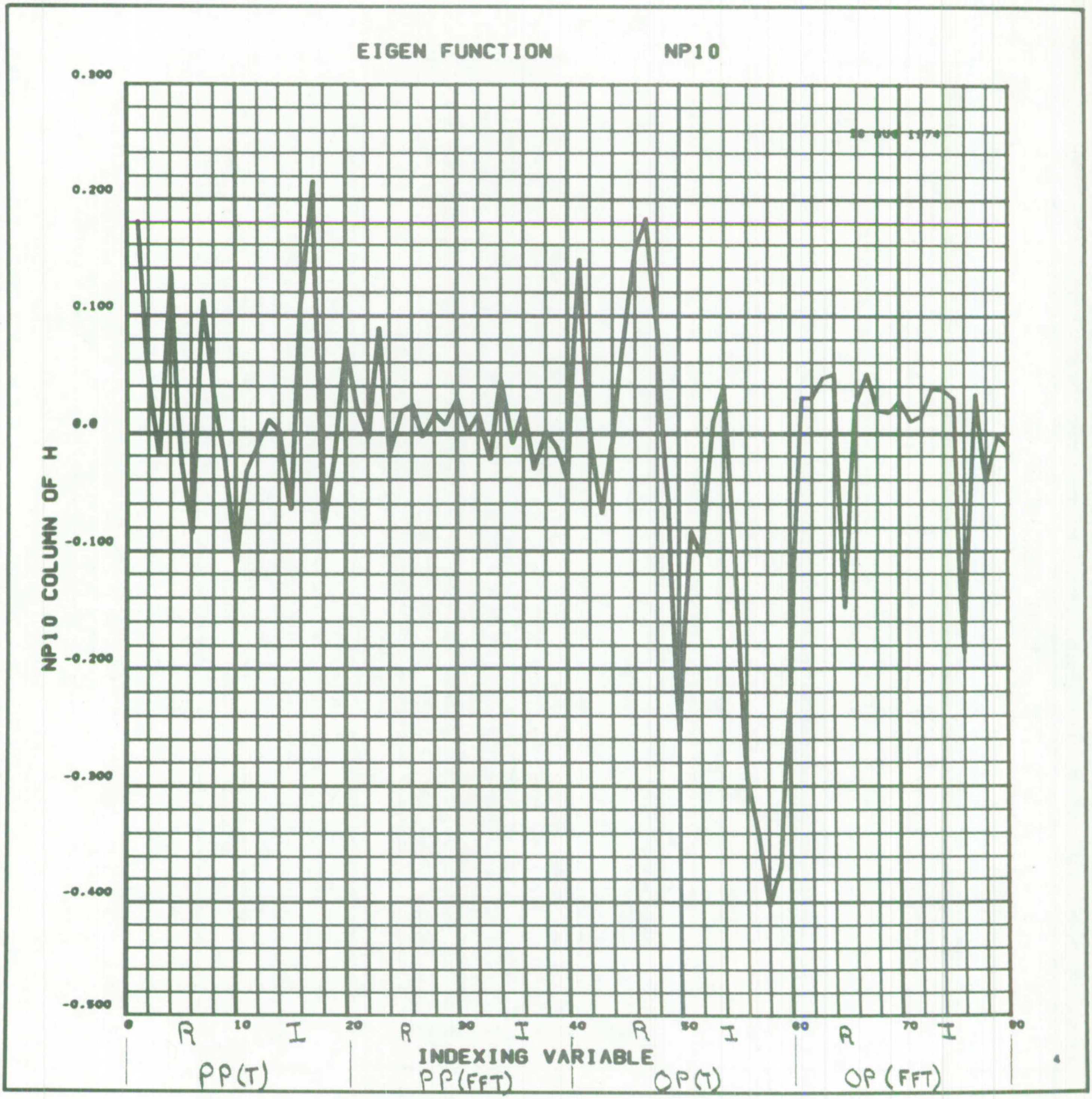


FIGURE 3.16 - ELEVENTH OPTIMAL FUNCTION REFERENCE BASE

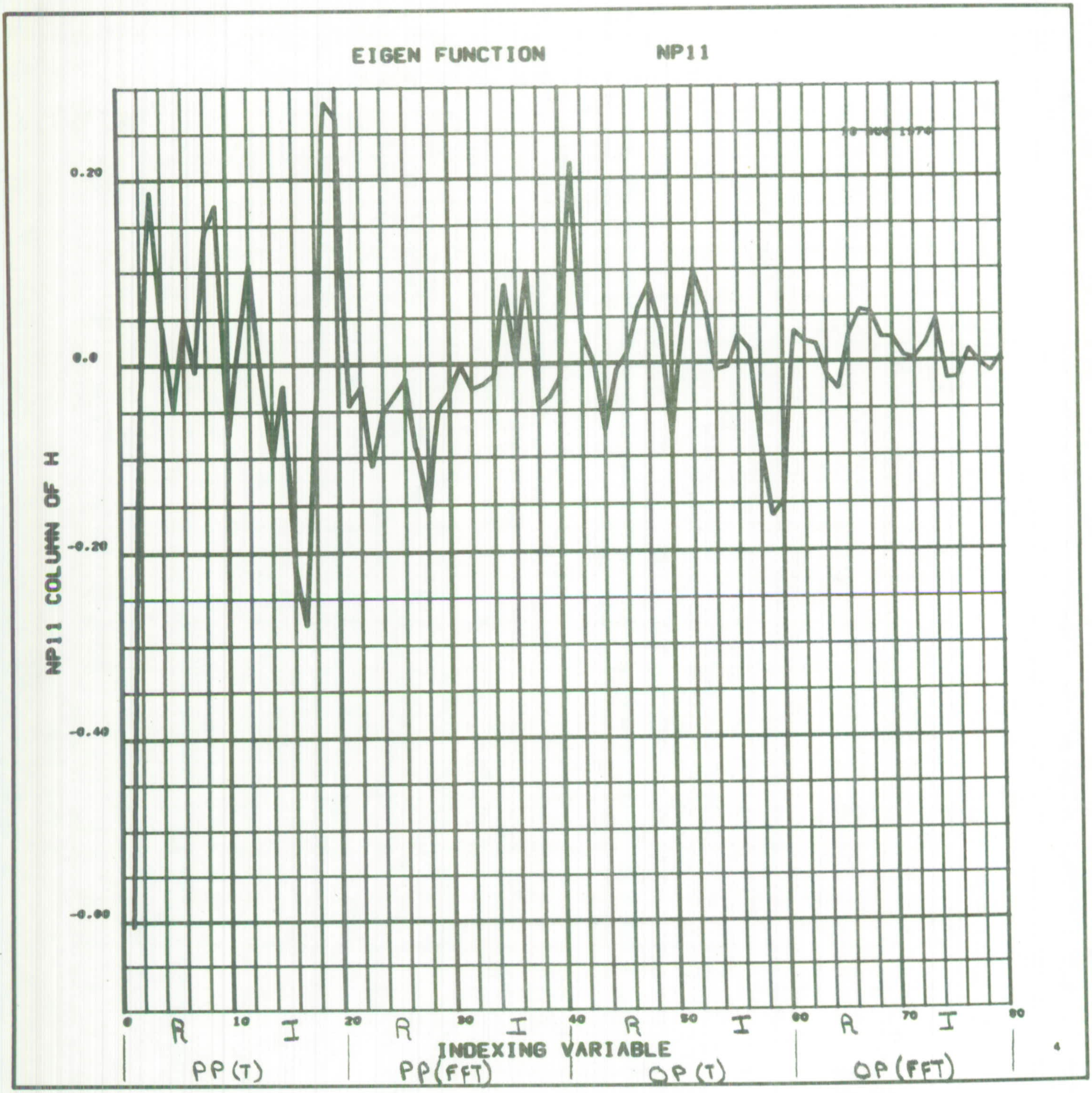


FIGURE 3.17 - TWELVETH OPTIMAL FUNCTION REFERENCE BASE

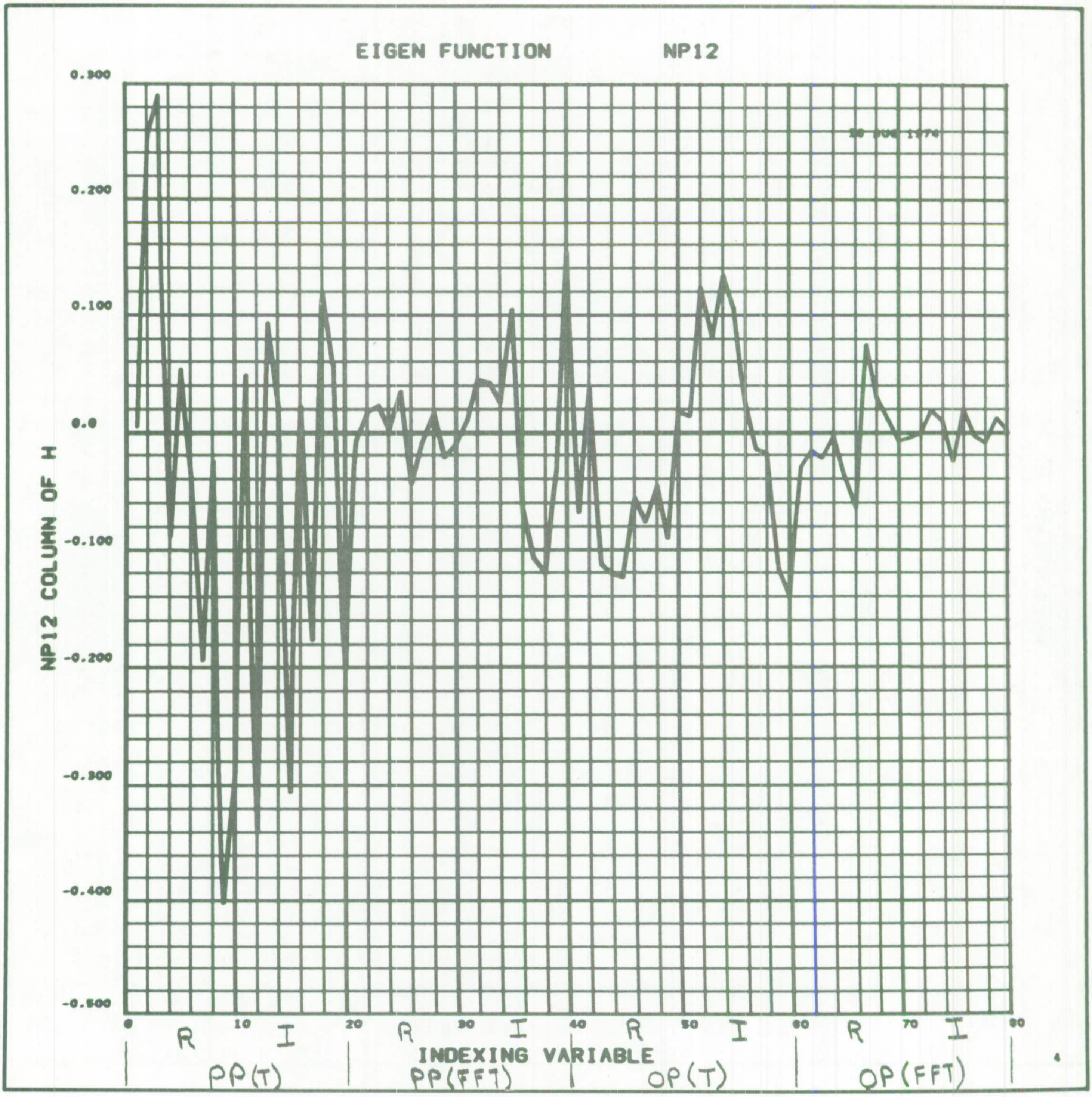


FIGURE 3.18 - THIRTEENTH OPTIMAL FUNCTION REFERENCE BASE

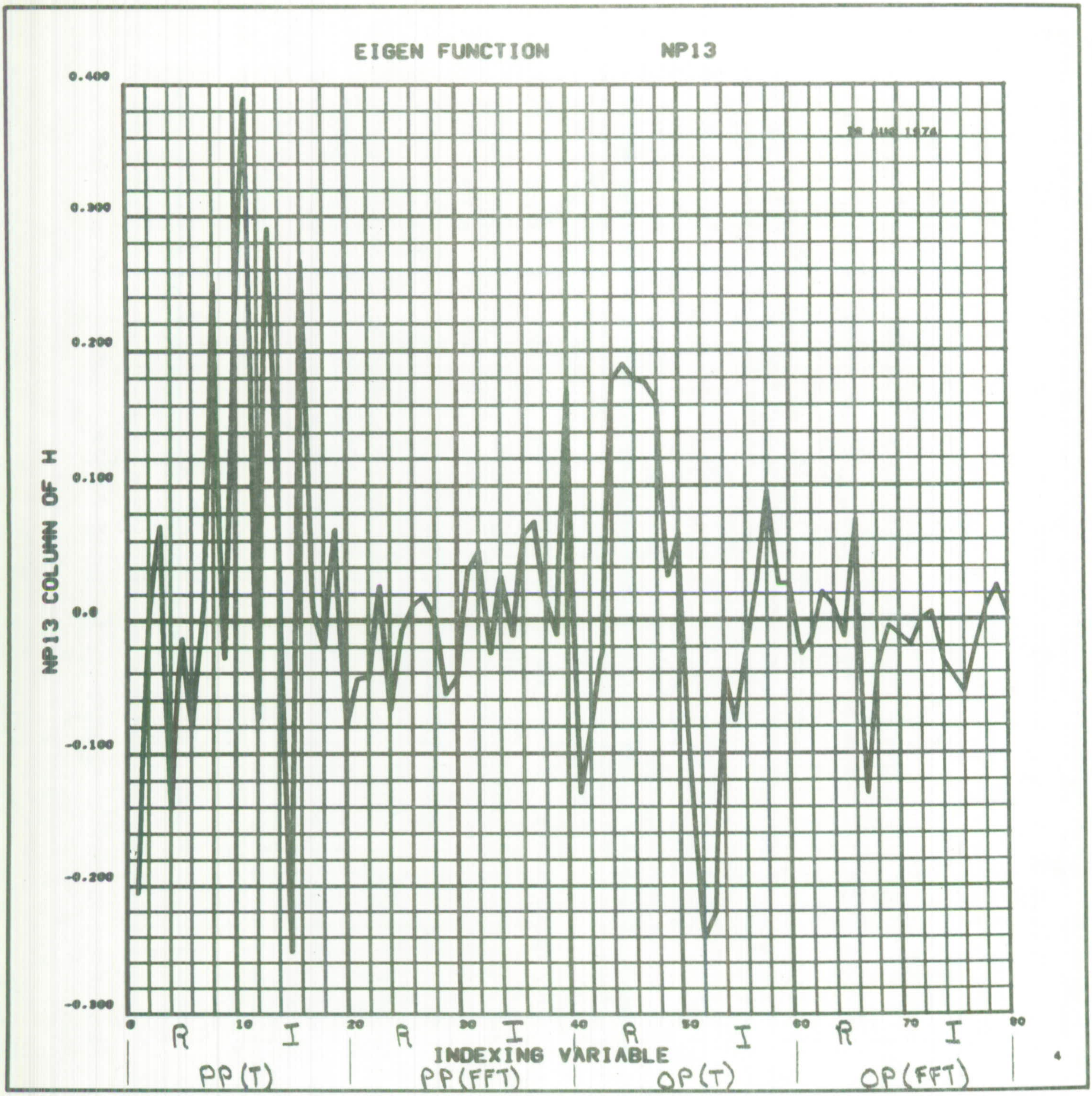


FIGURE 3.19 - FOURTEENTH OPTIMAL FUNCTION REFERENCE BASE

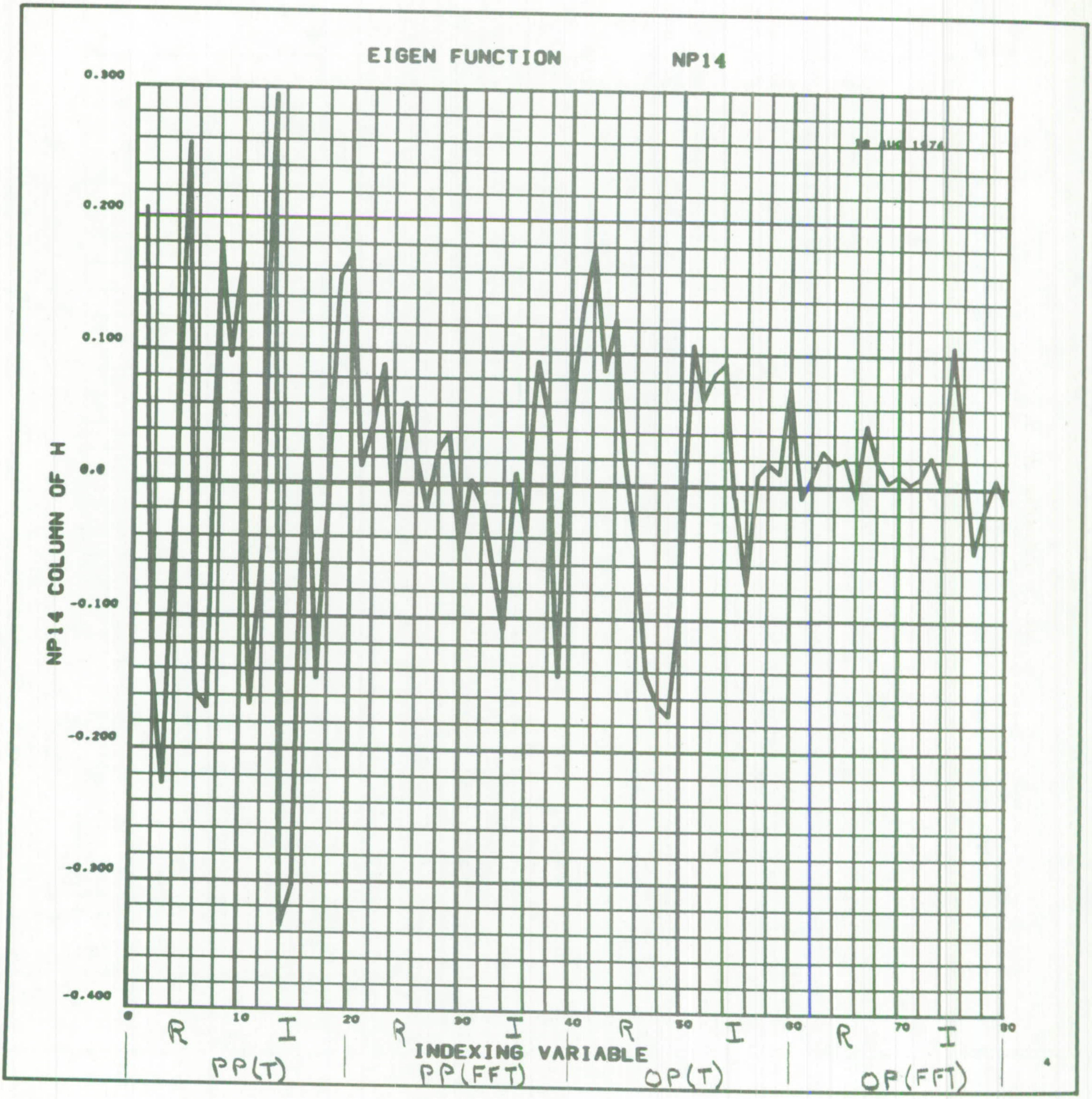
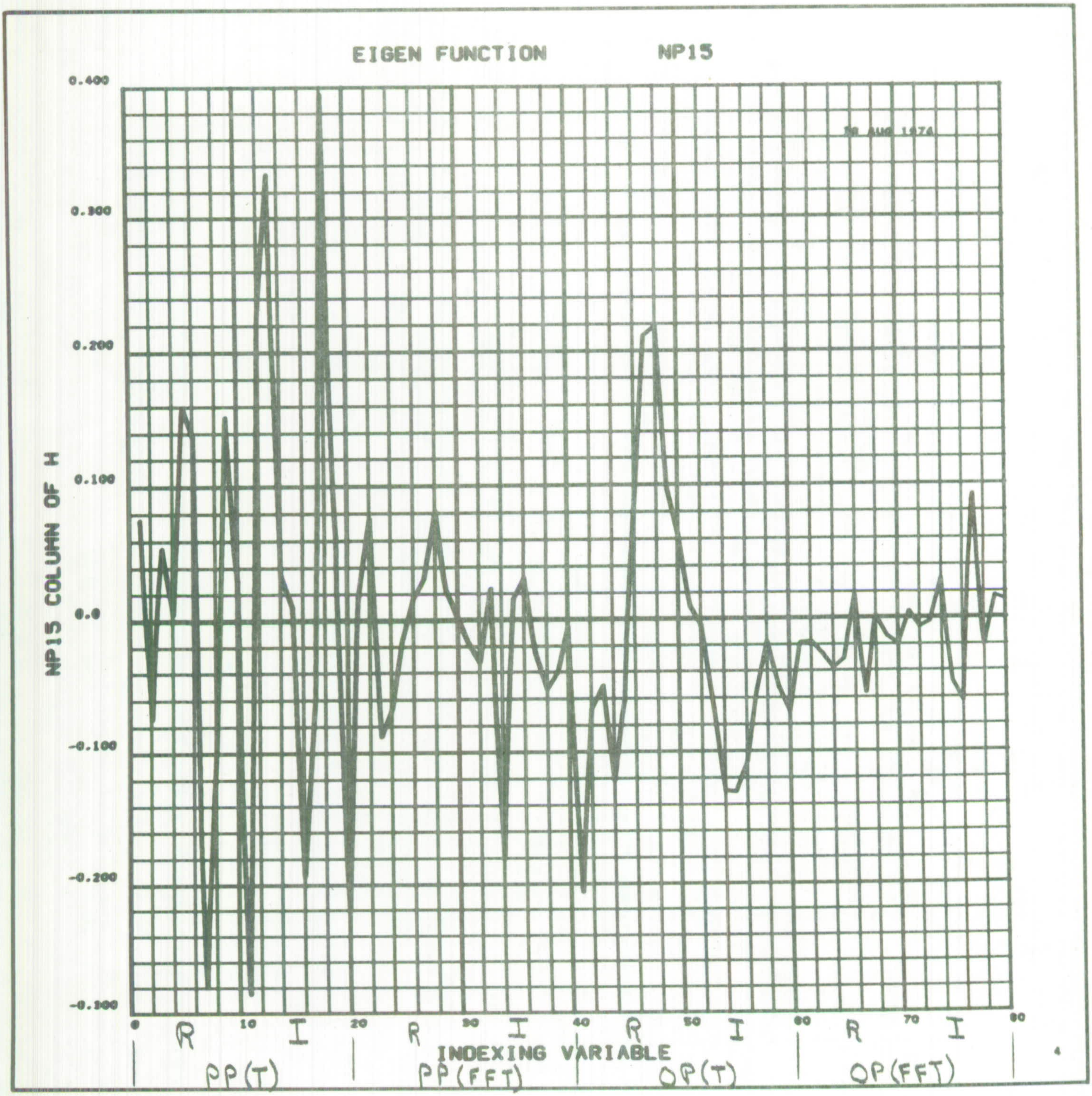


FIGURE 3.20 - FIFTEENTH OPTIMAL FUNCTION REFERENCE BASE



in the time domain than in the frequency domain.

3.2 Deletion of Targets to Demonstrate Generality of Bases

The easiest way to demonstrate the generality of a base for representing different target types then used in the learning data is to use that base to represent additional targets. However, for the present study, all of the available targets were used in the reference base; thus, this question was addressed by creating a new base in which one of the targets was deleted. The characteristic of this base and results obtained using it were used to show the insensitivity to a new target type. Two such bases were constructed, the first was the "3-Class-2 target" base in which one of the four 21X targets, target 214, was omitted from the learning data. The other base was a base constructed without the 101 target. The results of these comparison show that the reference base should be very insenitive to the addition of either additional 101 targets or additional 21X type targets.

Figure 3.21 shows the average of the 220 data histories not including the 214 target used to construct the first off reference optimal base. Comparison of Figure 3.3 and 3.21 shows only minor differences between these averages. The information energy curve for this base has not been presented because it is identical to the information energy curve for the reference base shown in Figure 3.4. This indicates that the bases are quite similar. The first four optimal functions of this base were also sufficiently similar to those presented in Figures 3.5 through 3.9 that the reader could not observe the differences. Figure 3.22 shows the fifth optimal function of the "3-Class 2 target" base. Comparison of Figures 3.10 and 3.22 shows that these two optimal functions are still quite similar although careful examination will reveal a few minor differences. Figure 3.23 presents the sixth optimal function for the "3-Class 2 target" base which can be compared with the sixth optimal function of the reference base presented in Figure 3.11. Although, initially these two optimal functions may appear quite different, they are very similar. The difference apparent to the eye is that Figure 3.23 is essentially the mirror image of the curve presented in Figure 3.11. This is merely a change in the sign of the coefficient corresponding to this optimal function or a change

FIGURE 3.21 - AVERAGE OF ALL 220 CASES USED TO CONSTRUCT THE THREE CLASS II TARGET BASE

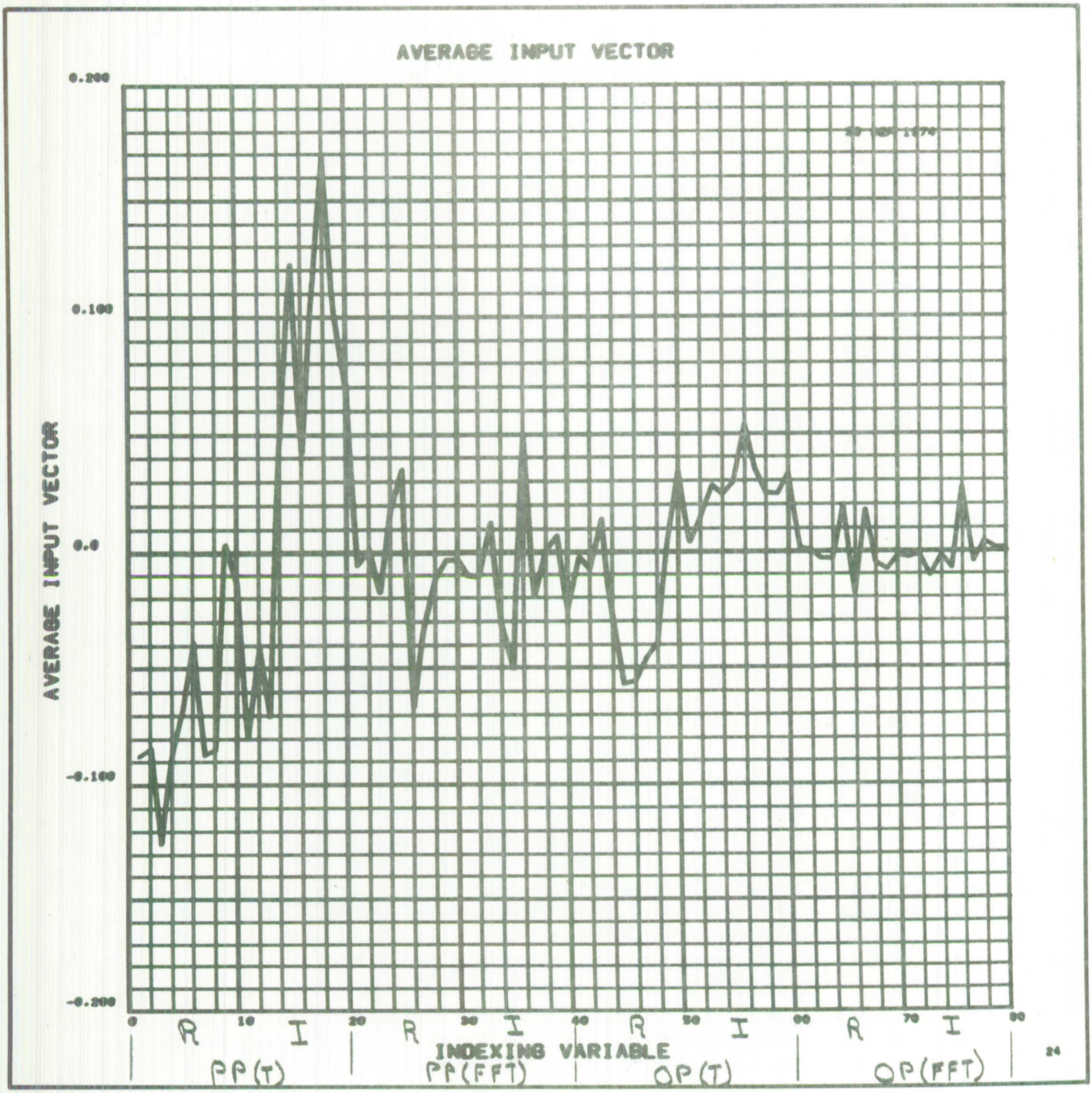


FIGURE 3.22 - FIFTH OPTIMAL FUNCTION THREE CLASS II TARGET BASE

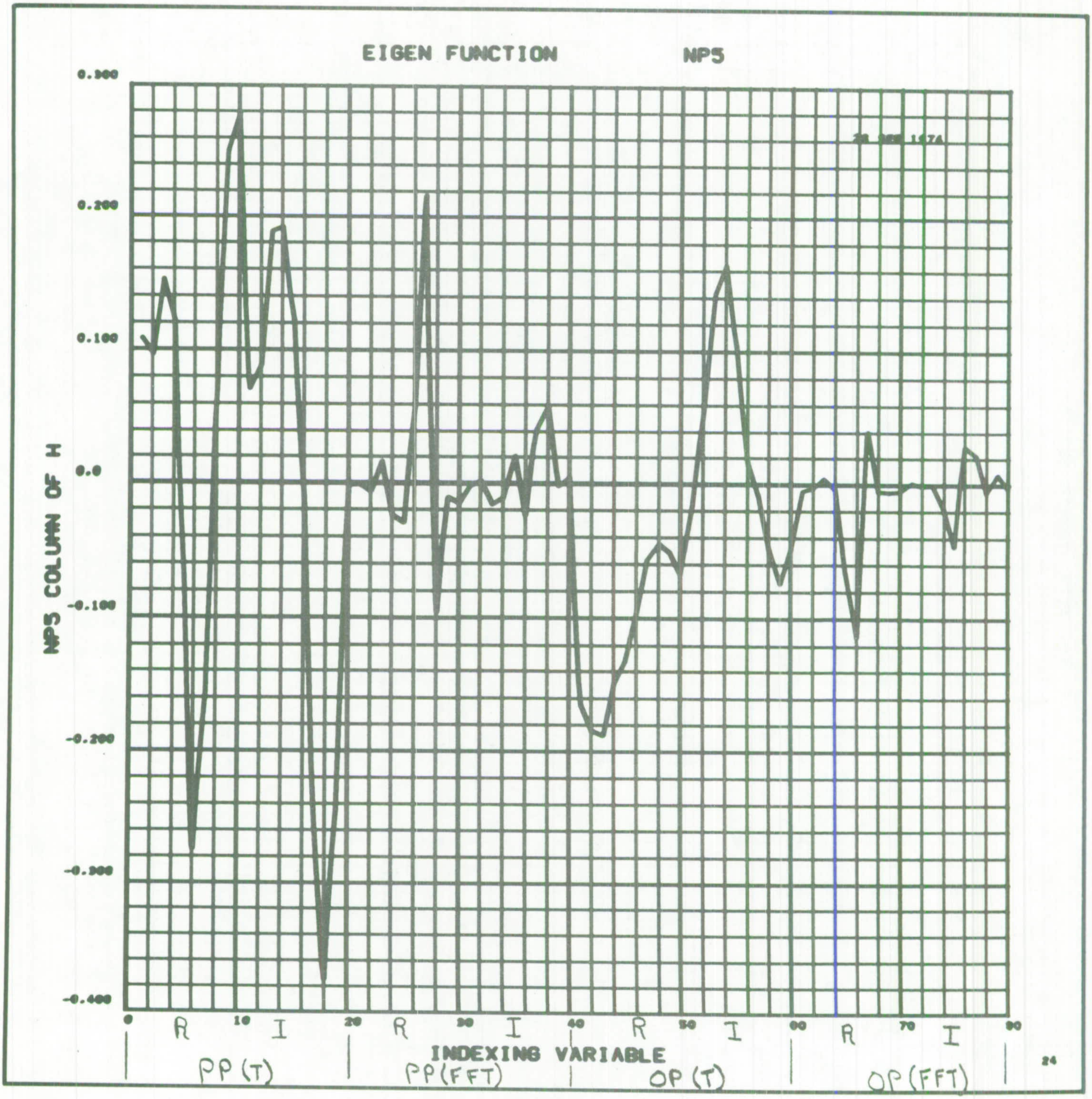
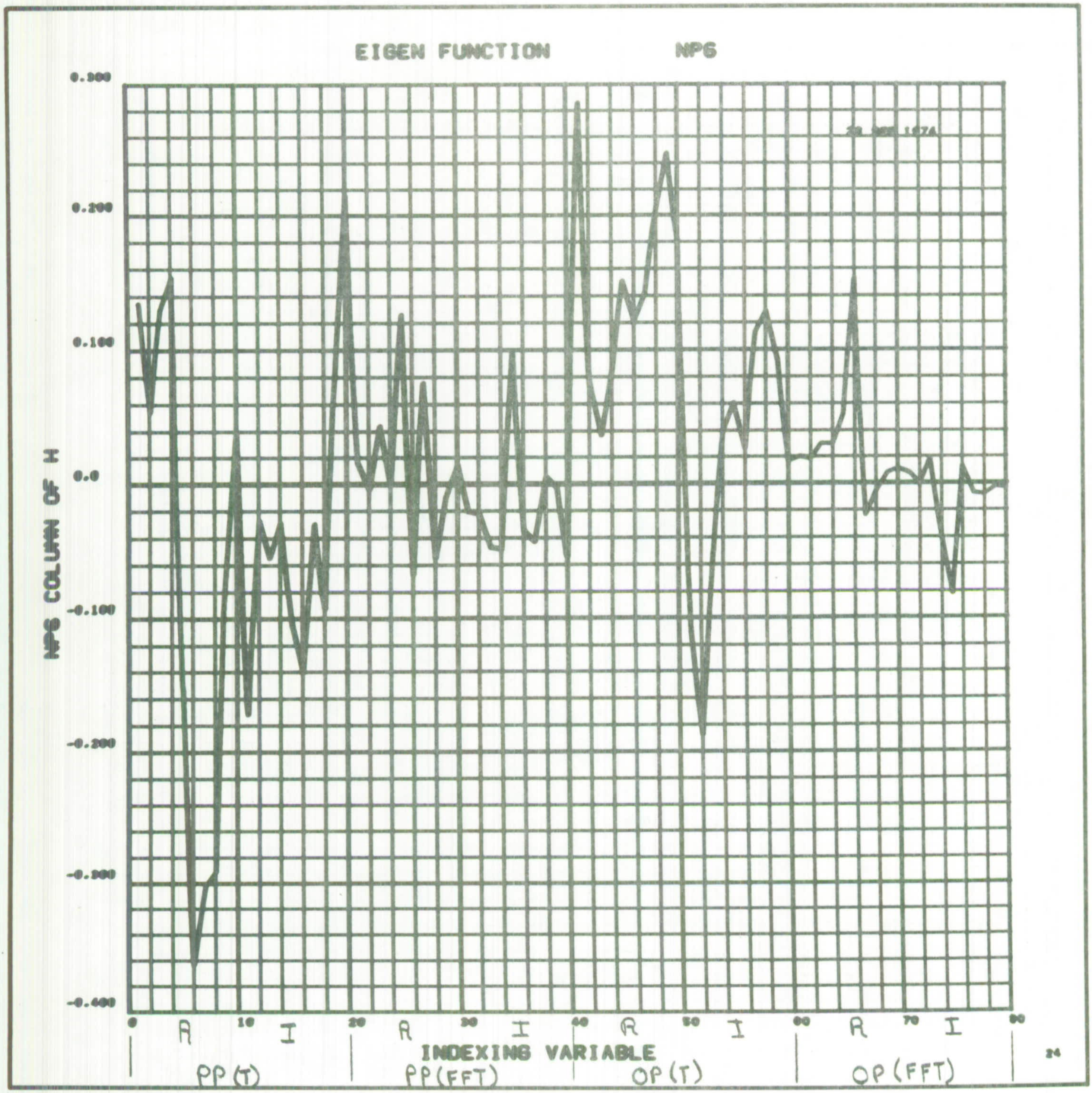


FIGURE 3.23 - SIXTH OPTIMAL FUNCTION THREE CLASS II TARGET BASE



in the definition of the negative or positive direction in the optimal space. Since this sign is cancelled by the corresponding sign on the coefficient, it is not significant. Since Figure 3.23 is the mirror image of Figure 3.11 there is still considerable similarity at the sixth optimal function. Figure 3.24 presents the seventh optimal function. Comparison of this with the seventh optimal function for the reference base presented in Figure 3.12 shows that there are very significant differences, however, if one compares Figure 3.24 with the eighth optimal function for the reference base presented in Figure 3.13, we again see similarities. The eighth optimal function for the "3-Class 2 target" base presented in Figure 3.25 is just a mirror image of the seventh optimal function for the referenced base. Thus, the seventh and eighth optimal functions have exchanged roles between these two bases. This is a relatively common phenomena when two ADAPT bases are constructed from similar data.

The information contained in the first 14 optimal functions is essentially the same for both the reference base and the "3-Class 2 target" base. Figure 3.27 presents the 14th optimal function which can be compared with the corresponding 14th optimal function of the reference base presented in Figure 3.19.

Thus, Figure 3.4 shows that for the most highly correlated 90% of the variation in the data set, the "3-Class 2 target" and the reference base are very much the same. Since the analysis carried out in the present study only used 15 dimensions the results obtained in the present study should not be sensitive to which of these two bases are used. This implies that if one were to add additional targets of the 21X class, containing similar variation to that existing between the four members of the 21X class used in the present study, there should be very little difference in the optimal base for representing the ensemble of the original plus the new data.

Next consider effect of using additional 101 class targets. In most circumstances, it would be impossible to obtain an answer to this question with only one 101 class target. However, the characteristic of the bases obtained from this data also allows one to demonstrate the insensitivity of the base to the 101 class. The scatter plots showed that a very small amount of the variation in the data is due to

FIGURE 3.24 - SEVENTH OPTIMAL FUNCTION THREE CLASS II TARGET BASE

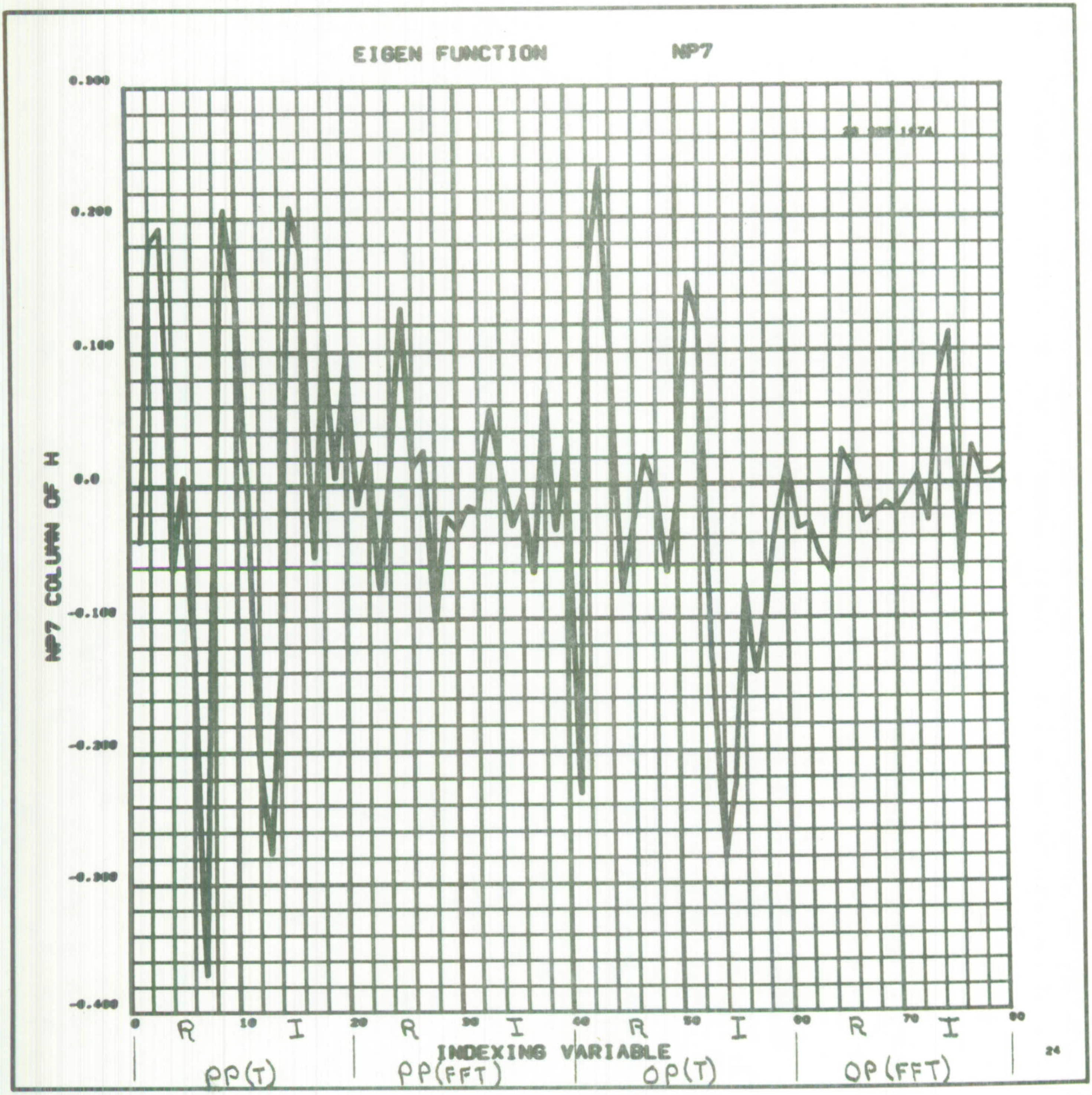


FIGURE 3.25 - EIGHTH OPTIMAL FUNCTION THREE CLASS II TARGET BASE

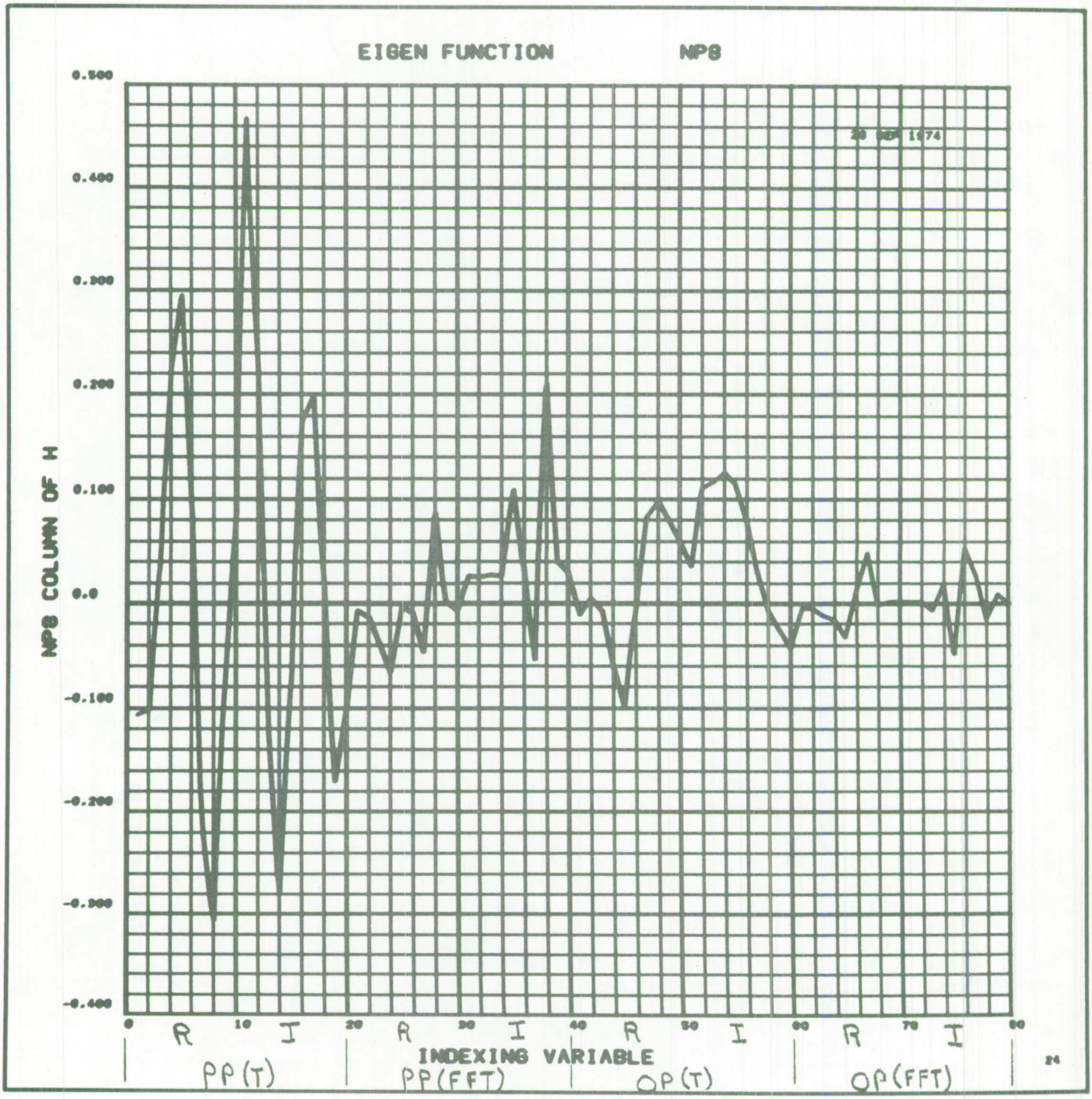


FIGURE 3.26 - THIRTEENTH OPTIMAL FUNCTION THREE CLASS II TARGET BASE

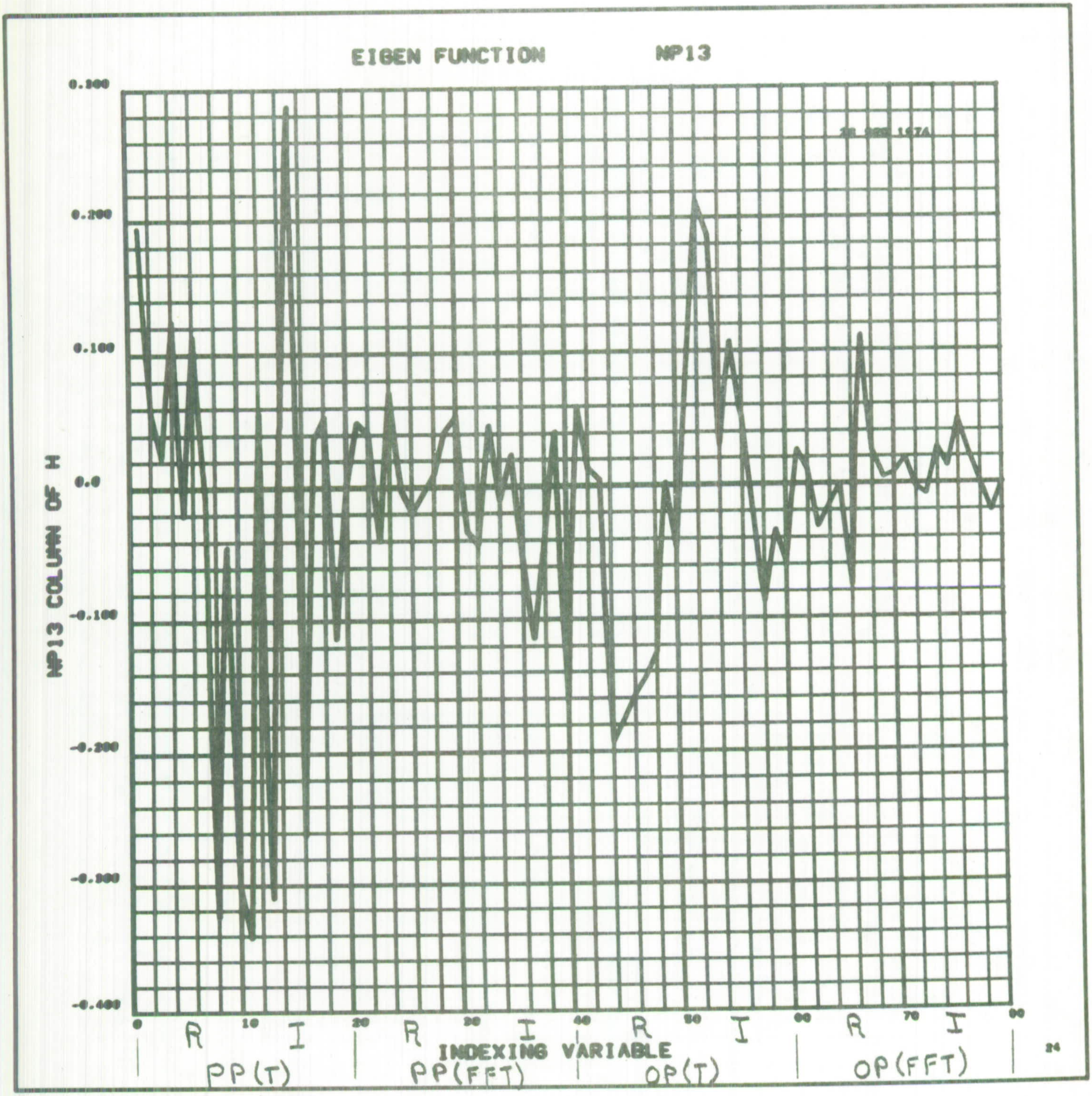
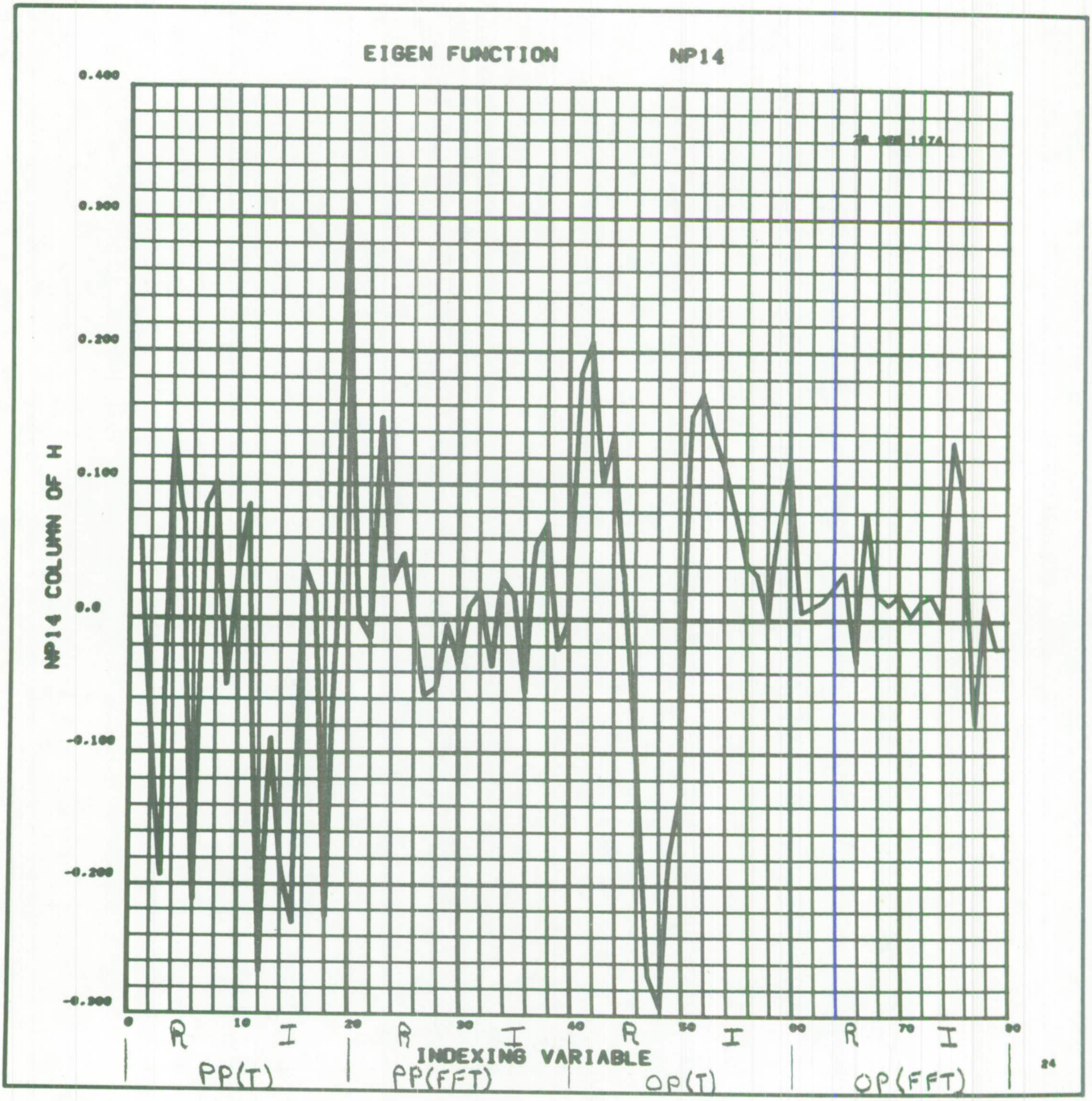


FIGURE 3.27 - FOURTEENTH OPTIMAL FUNCTION THREE CLASS II TARGET BASE



the 101 class. Thus, one might hypothesize that the base is defined by the 21X class. In other words, the variation in the 21X class is sufficient to cover all the possible variations in the 101 class and a base derived from only the 21X class is the same as a base derived from both the 21X and 101 class.

To verify this hypothesis, a new base was constructed using only the 21X targets. Figure 3.28 shows the average of all of the 21X targets. Comparing Figures 3.28 and Figure 3.3 we see that the major difference is one of scale, this implies that the 101 targets have exactly the same average as the 21X targets except for a smaller amplitude or that the 101 targets have essentially zero for their average input vector. The base constructed using only the 21X data was so similar to the reference base that no significant differences could be seen in either the information energy curve or the first 15 optimal functions. To illustrate this, the fifteenth optimal function for the Class 21X only base is presented in Figure 3.29. The reader may easily verify the similarity of these two optimal functions by comparing this function with the fifteenth optimal function for the reference base presented in Figure 3.20.

3.3 Low Variation Sub Space

Figure 3.7 showed that the 101 targets all fell in a very small portion of the first two dimensions of the optimal space. Examination of the scatter plots up to the fifteenth dimension shows that this is true in general and that the 101 target is restricted to a very small region of the optimal space. This implies that a nearest neighbor discriminant to this region is quite effective in rejecting a large percentage of the 21X targets. However, some of the 21X targets also fall in the same region as the 101 targets. Thus, a base was developed which only encompassed this portion of the space. This base was constructed using only the 101 targets and the 21X targets which fell in the same general region of the optimal space. This base has been designated the low variation base and the average of all of the cases used for this base are presented in Figure 3.30. Comparison with the average for the reference base shown in Figure 3.3 shows that there are significant differences both

FIGURE 3.28 - AVERAGE OF ALL 160 CLASS II DATA HISTORIES

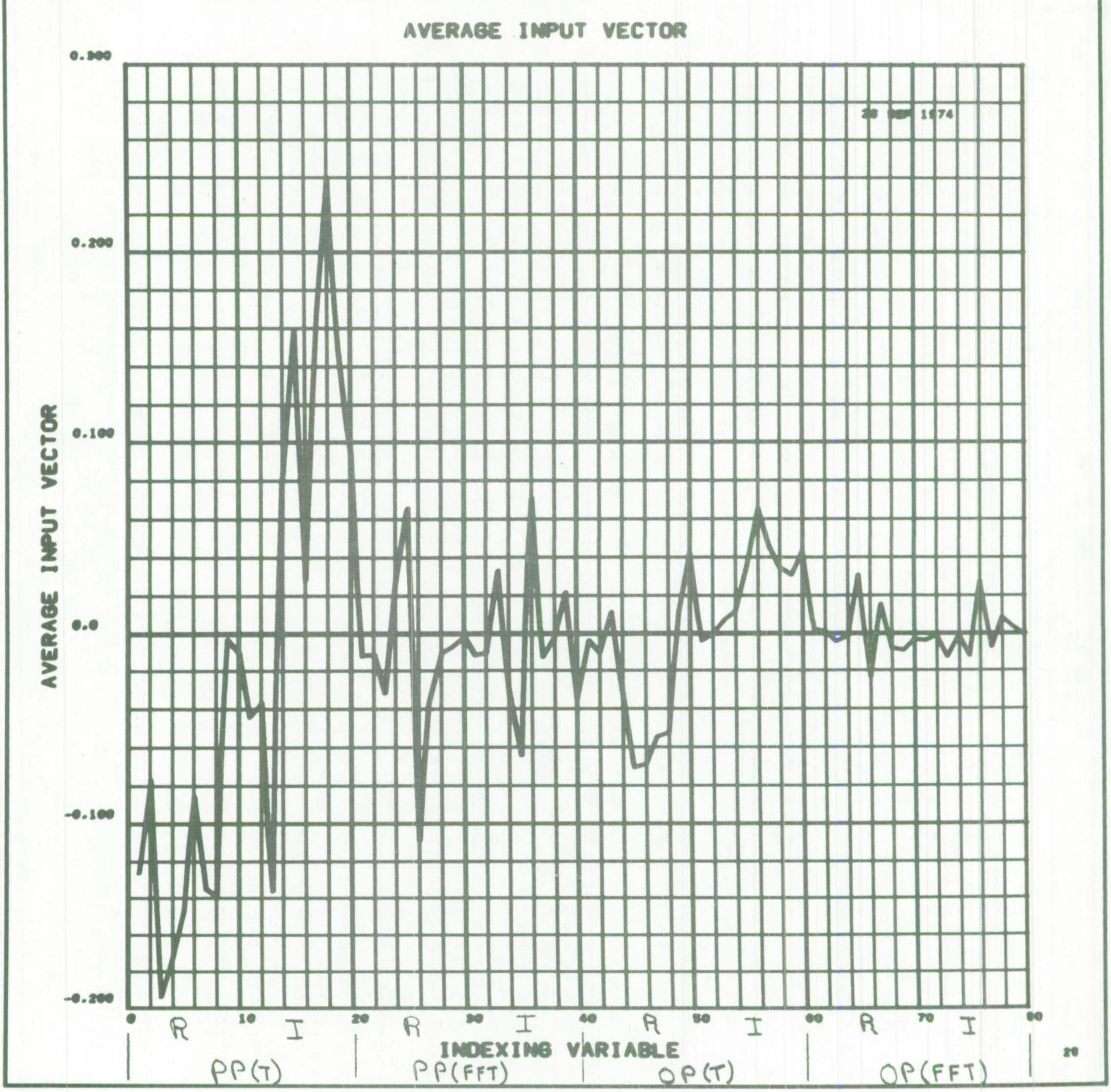


FIGURE 3.29 - FIFTEENTH OPTIMAL FUNCTION FOR CLASS II ONLY BASE

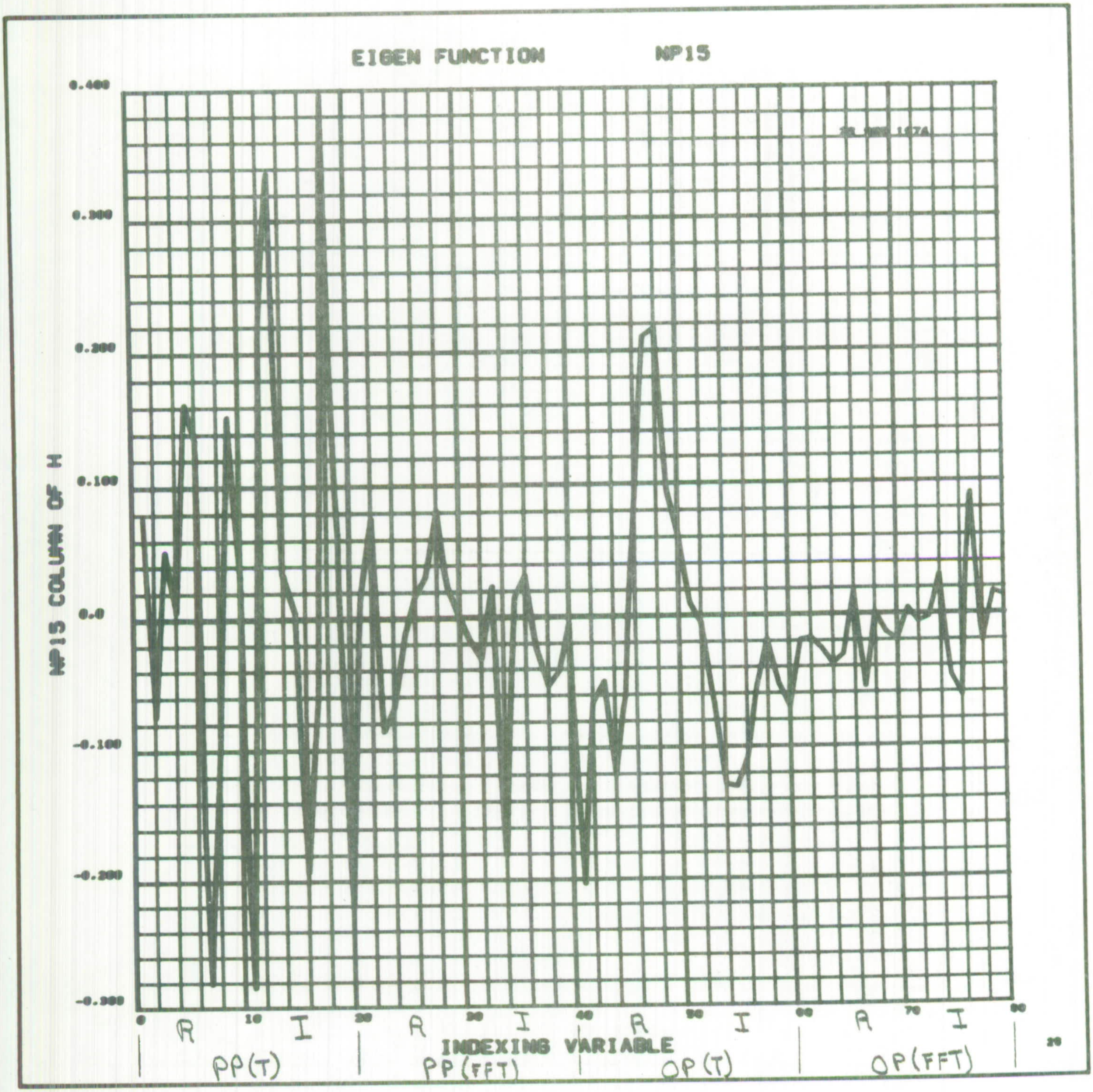
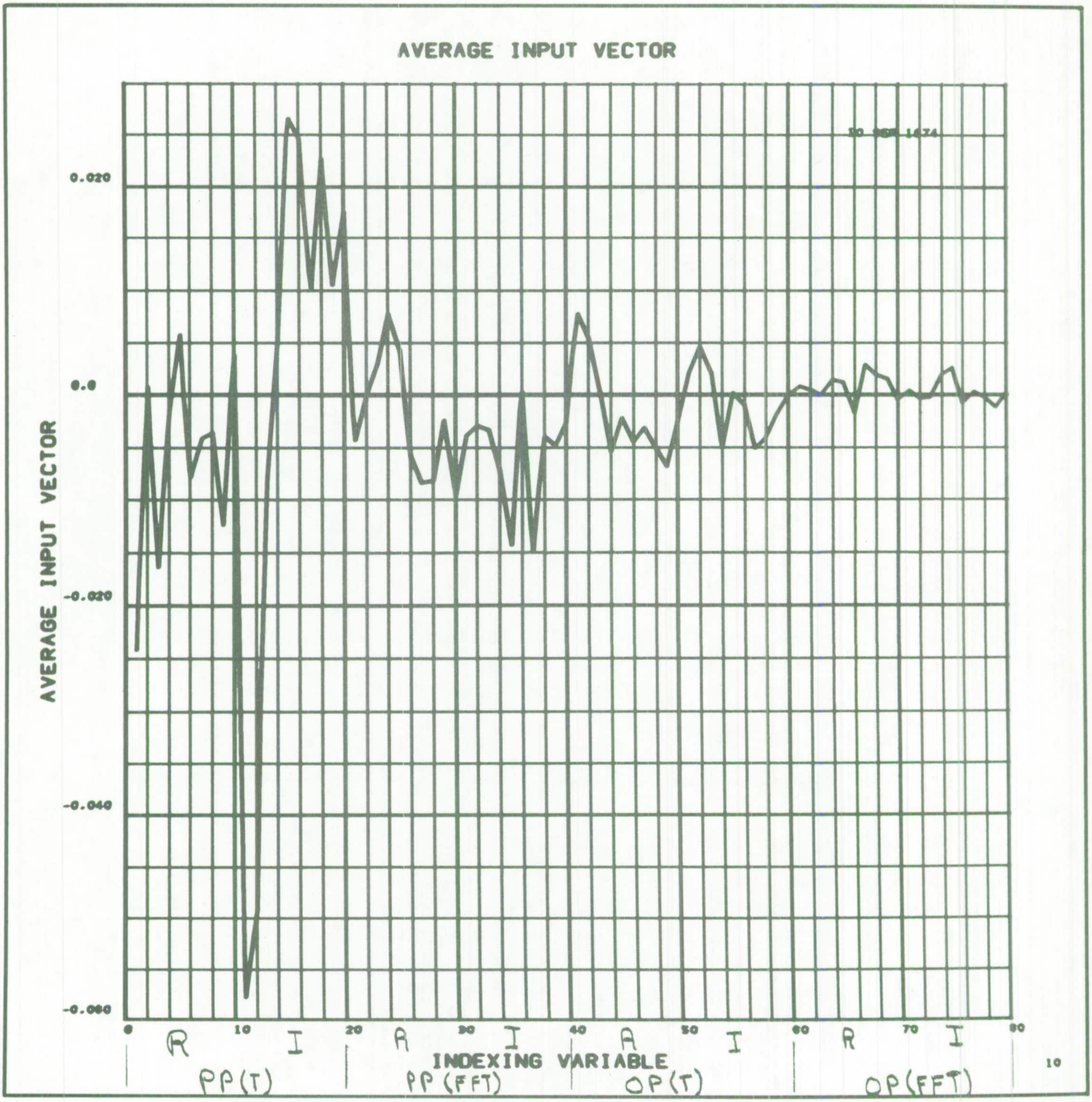


FIGURE 3.30 - AVERAGE OF ALL 116 LOW VARIATION CASES USED FOR LOW VARIATION BASE



in shape and in scale. Note, that the scale on Figure 3.30 is almost an order of magnitude smaller than that on Figure 3.3. This suggests that these cases in addition to possessing considerably less variation also have a lower average value.

Figure 3.31 shows the explained variation as a function of the number of dimensions used for this base. Since this base does not have to explain as much variation as the reference base, it is not surprising that the first term explains a considerably larger percentage of the variation then was the case for the reference base. Similarly, the percent of information contained in any given number of terms tends to be larger than for the reference base. The first two optimal functions for the low variation base are presented in Figures 3.32 and 3.33. Comparison with the corresponding optimal functions for the reference base shows that the low variation base appears to have a more noise like structure. The scatter plot of the coefficients of the first two terms for the generalized Fourier series expansion of the cases used to make this base is presented in Figure 3.34. The low variation base is made utilizing 16 members of the 21X class and the entire 101 class. Examination of Figure 3.34 shows that the 101 class still groups in a considerably tighter group which in this case only contains one of the 16 21X members. Thus, the low variation base has not changed the general character of the classification problem which existed with the reference base.

3.4 Bases Illustrating Effect of Reducing Radar System Complexity

Bases were developed to investigate the effect of reducing the radar systems complexity, and therefore the information content of the radar signature, on the classification schemes. The three bases developed were: 1) a base using only the principal polarization, 2) a base which only used information in the time domain and omitted all information from the frequency domain and 3) a base which only used information from the principal polarization in the time domain. These three bases were designated the principal polarization, time and the principal polarization-time base, respectively.

FIGURE 3.31 - EFFECT OF DIMENSIONS ON AVAILABLE INFORMATION-
LOW VARIATION BASE

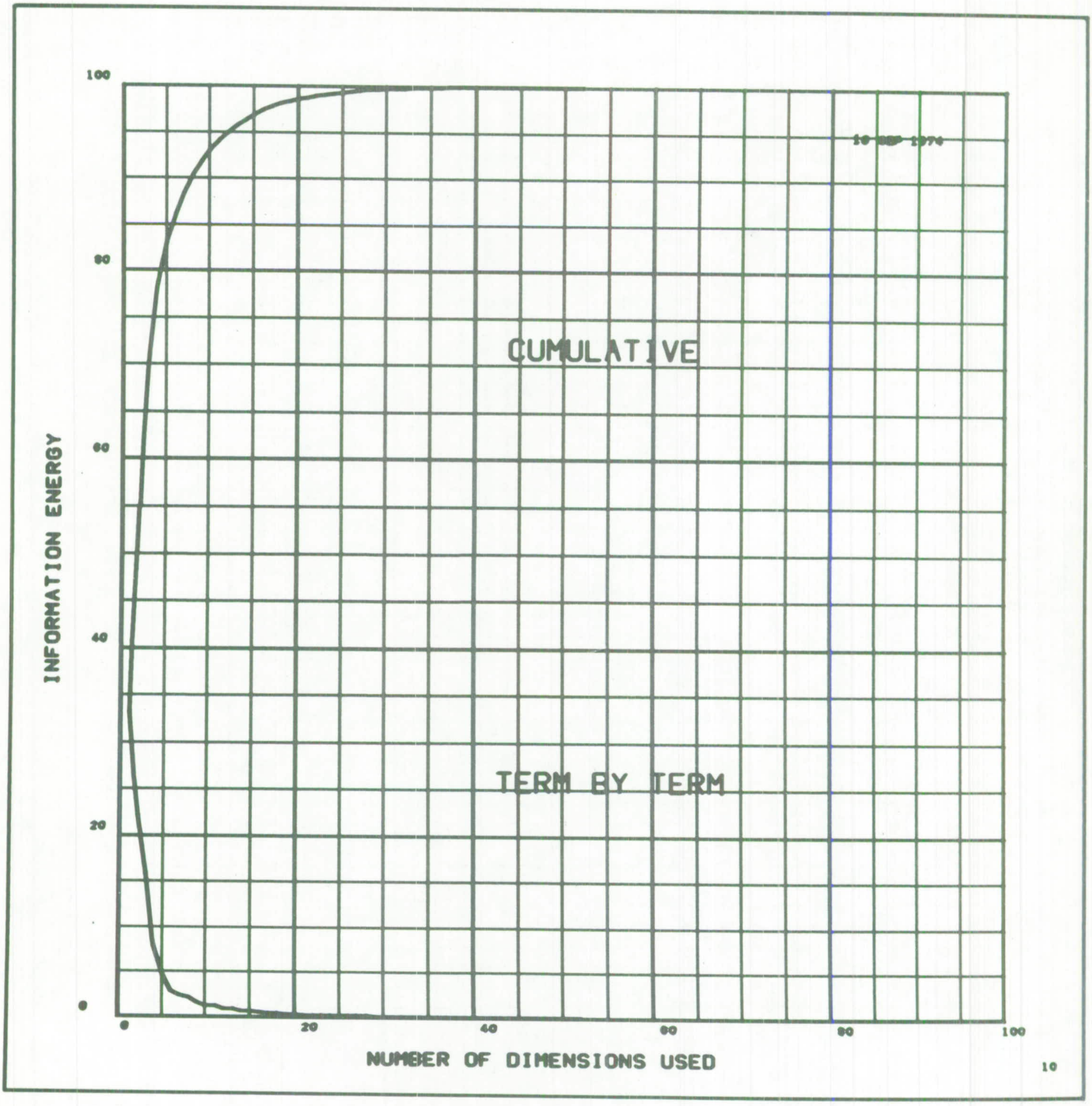


FIGURE 3.32 - FIRST OPTIMAL FUNCTION LOW VARIATION BASE

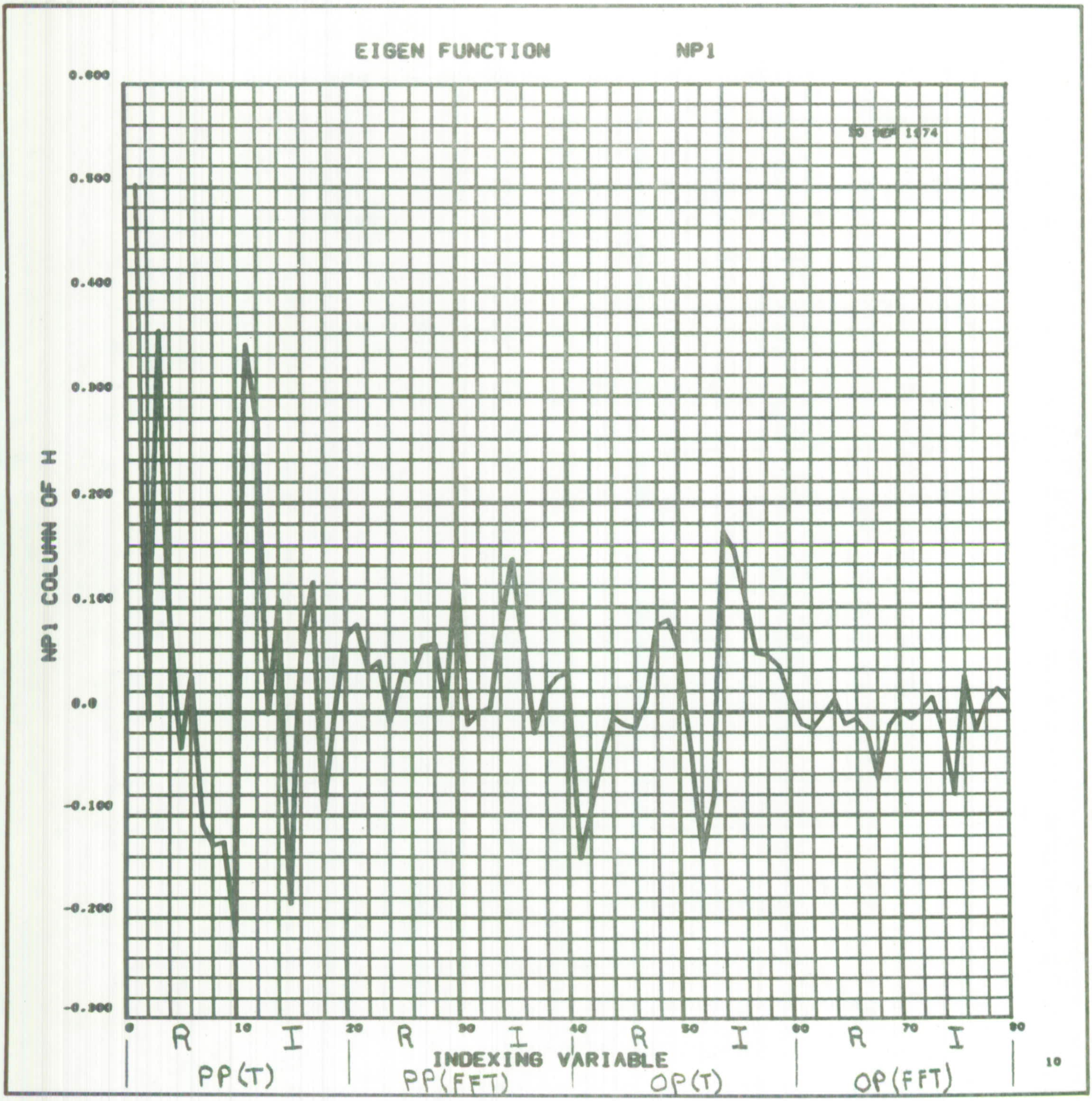


FIGURE 3.33 - SECOND OPTIMAL FUNCTION LOW VARIATION BASE

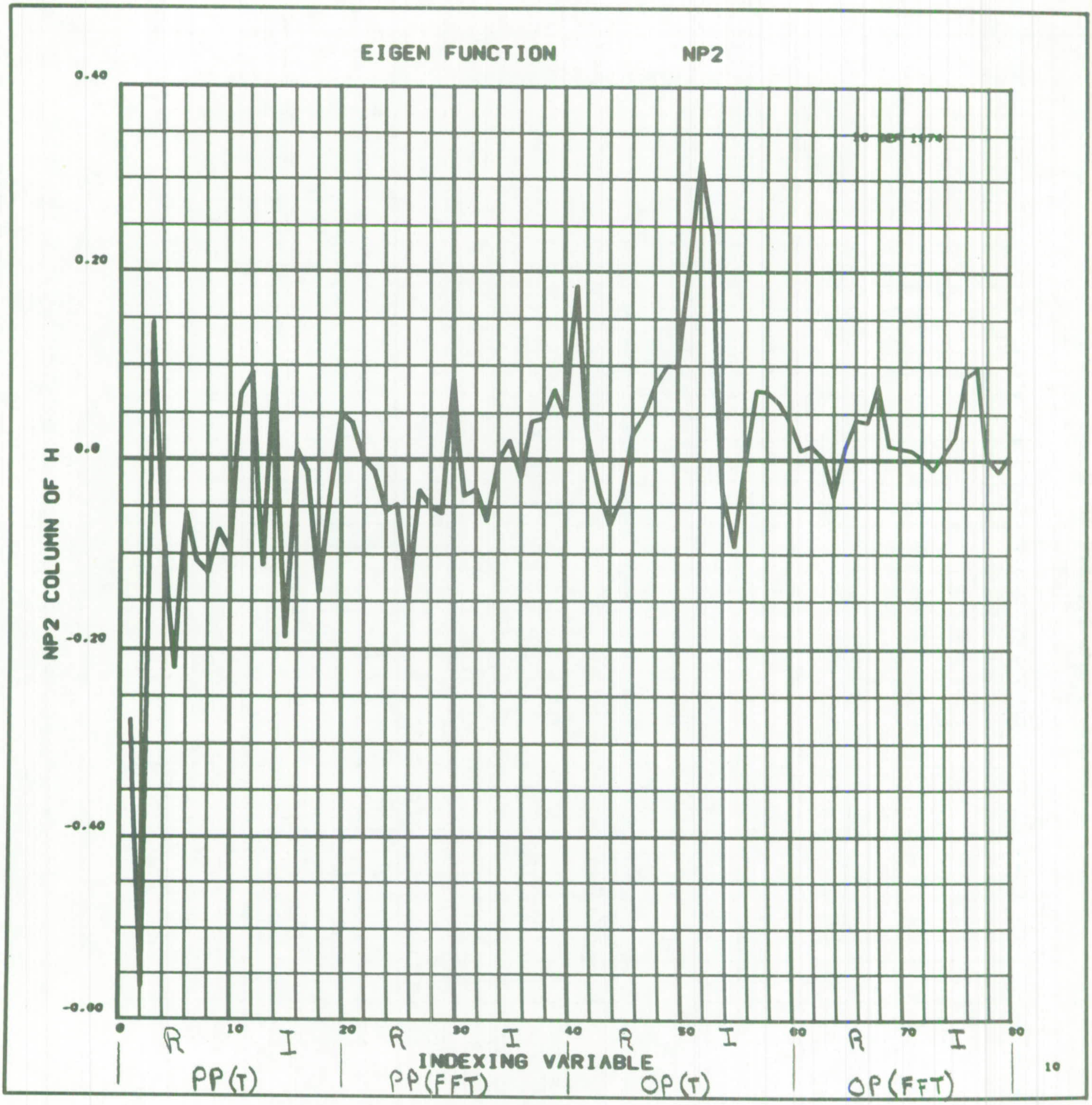
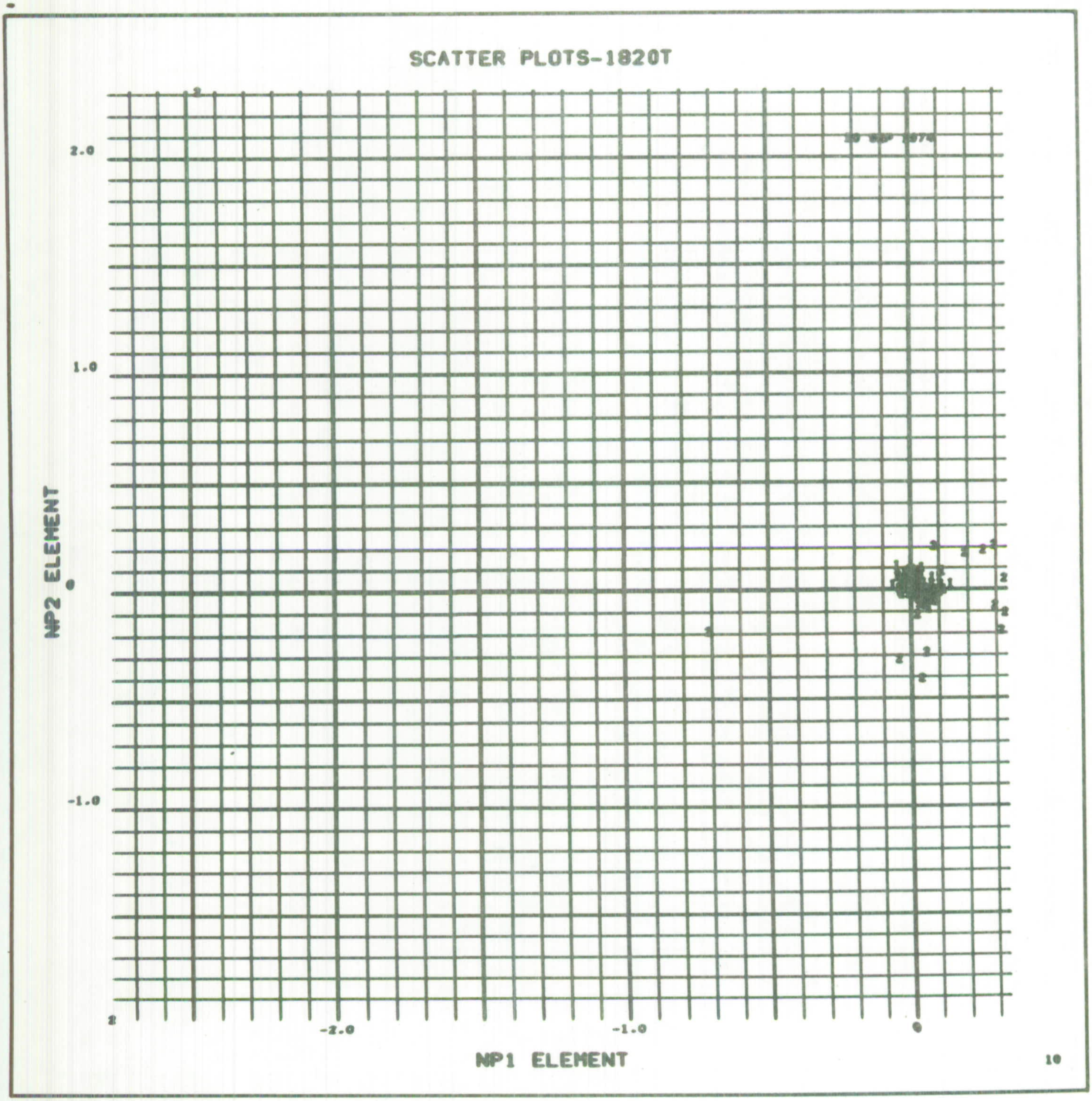


FIGURE 3.34 - SCATTER PLOT OF COEFFICIENTS OF FIRST AND SECOND TERMS IN OPTIMAL REPRESENTATION-LOW VARIATION BASE



The typical data histories and average input vectors for these bases are just the corresponding portions of Figures 3.1 through 3.3. For example, the typical data history for the principal polarization base is just that portion of Figure 3.1 including indexing variables 1 through 40. Similarly, the average input vector is that portion of Figure 3.3 including indexing variables 1 through 40.

Principal Polarization Base

The information energy or explained variation as a function of number of terms retained for the base utilizing only the information from the principal polarization of the radar return is presented in Figure 3.35. Since it is no longer necessary to represent the information from the opposite polarization, a given percentage of the variation should be explained with a smaller number of terms. This can be seen by comparison of Figure 3.35 with Figure 3.4. For the principal polarization base, the first term in the representation explains 25% of the information as contrasted to 20% for the reference base. The first 15 dimensions explain approximately 98% of information as compared with approximately 90% for the reference base.

The first optimal function for the principal polarization base is presented in Figure 3.36. Comparison of this figure with the corresponding points of the first optimal function of the reference base presented in Figure 3.5 shows that the first optimal function for the principal polarization base is themirror image of the corresponding points for the first function of the reference base. As pointed out previously, this mirror imaging is an insignificant difference and we conclude that the first optimal function represents the same physical phenomena even when only the principal polarization is used. The second optimal function presented in Figure 3.37 is also the mirror image of this portion of the second optimal function of the reference base. However, there are some minor differences in the detail structure especially in the frequency plane portion of the signature. Although the third optimal function for the principal polarization base is not similar to the first 40 points of the third optimal function of the reference base, it is very similar to the first 40 points of the fourth optimal function of the reference base. Thus, for the principal polarization base, the third optimal function contains the same physical information

FIGURE 3.35 - EFFECT OF DIMENSIONS ON AVAILABLE INFORMATION-
PRINCIPAL POLARIZATION BASE

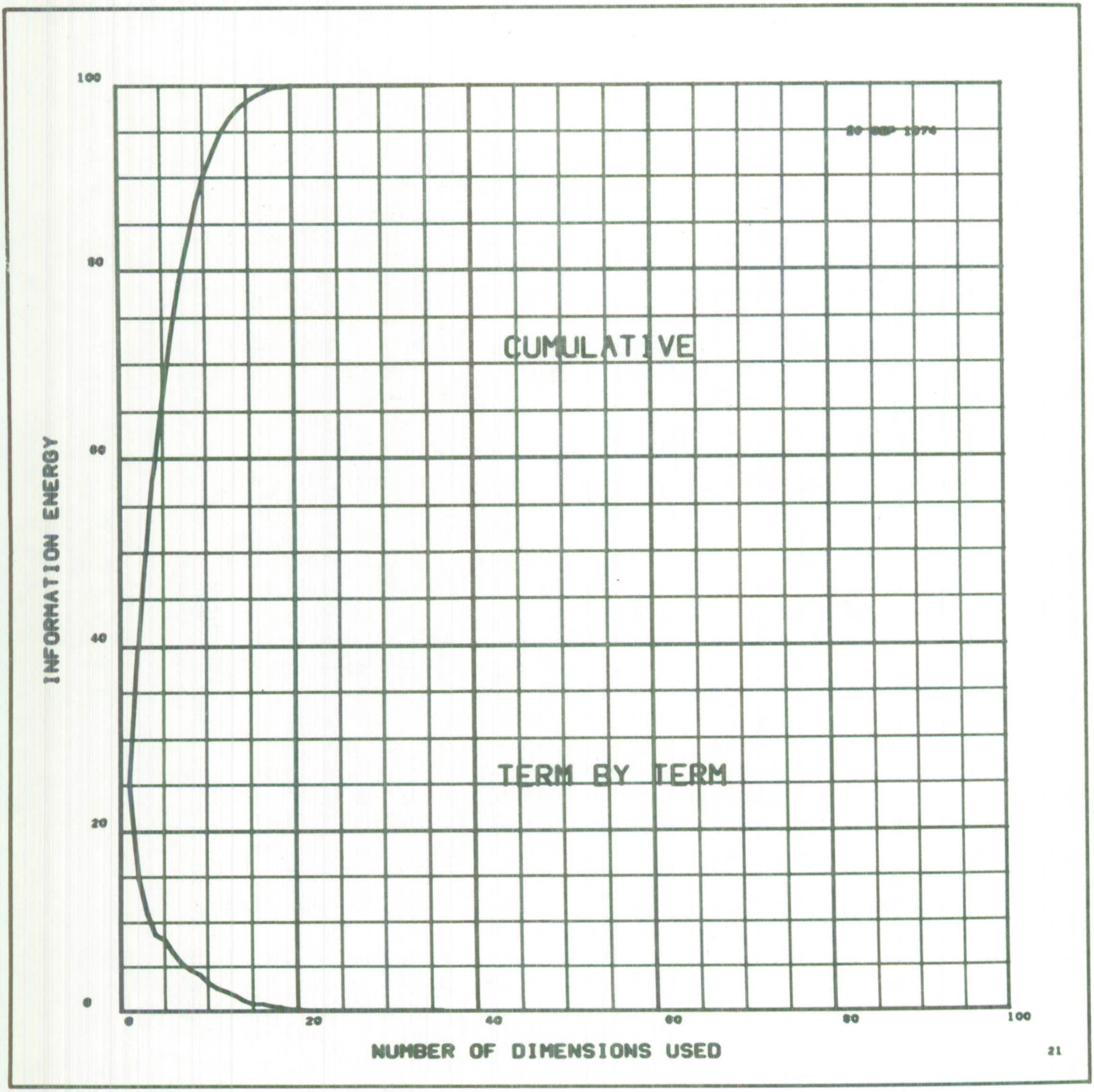


FIGURE 3.36 - FIRST OPTIMAL FUNCTION PRINCIPAL POLARIZATION BASE

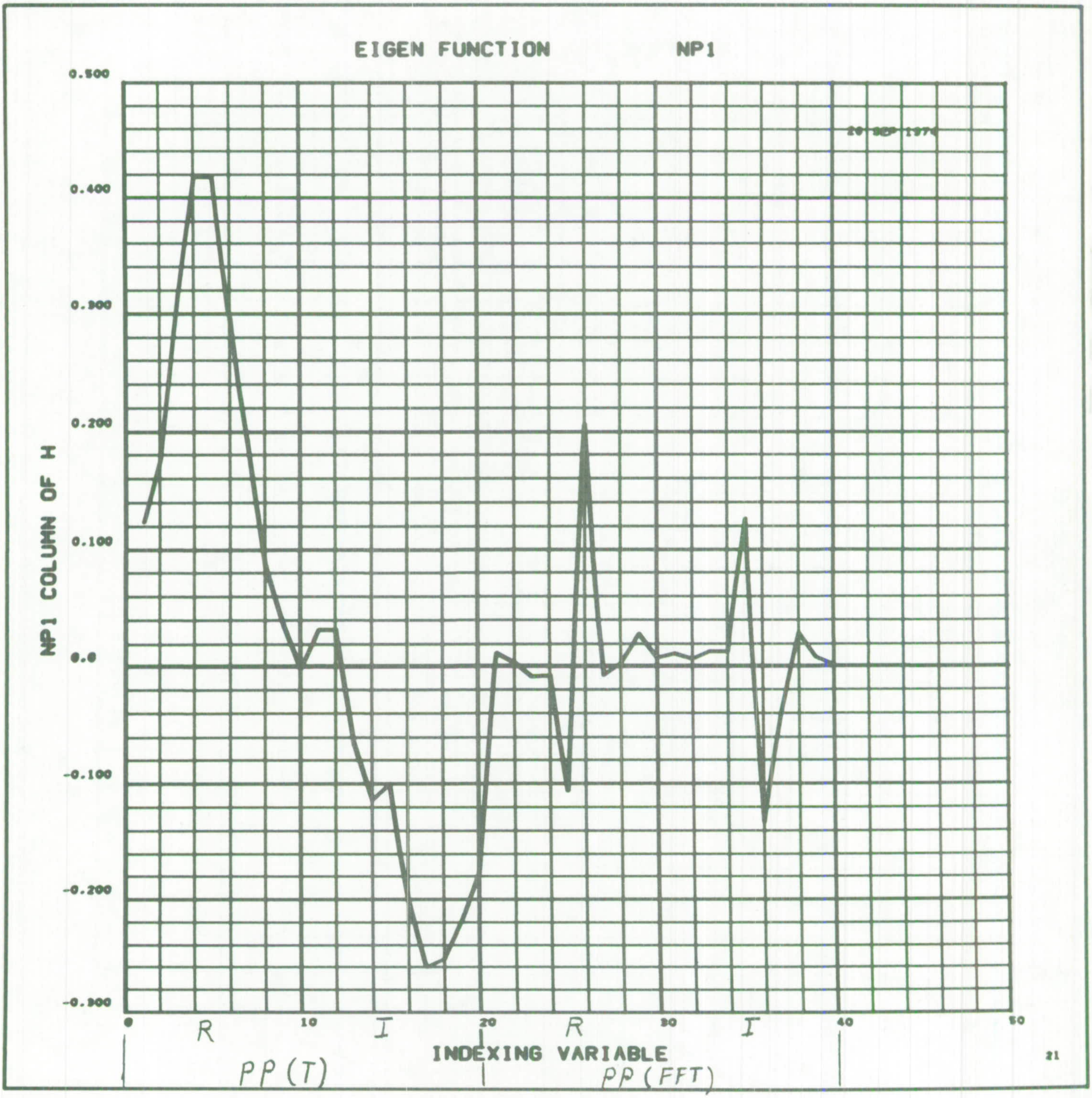


FIGURE 3.37 - SECOND OPTIMAL FUNCTION PRINCIPAL POLARIZATION BASE

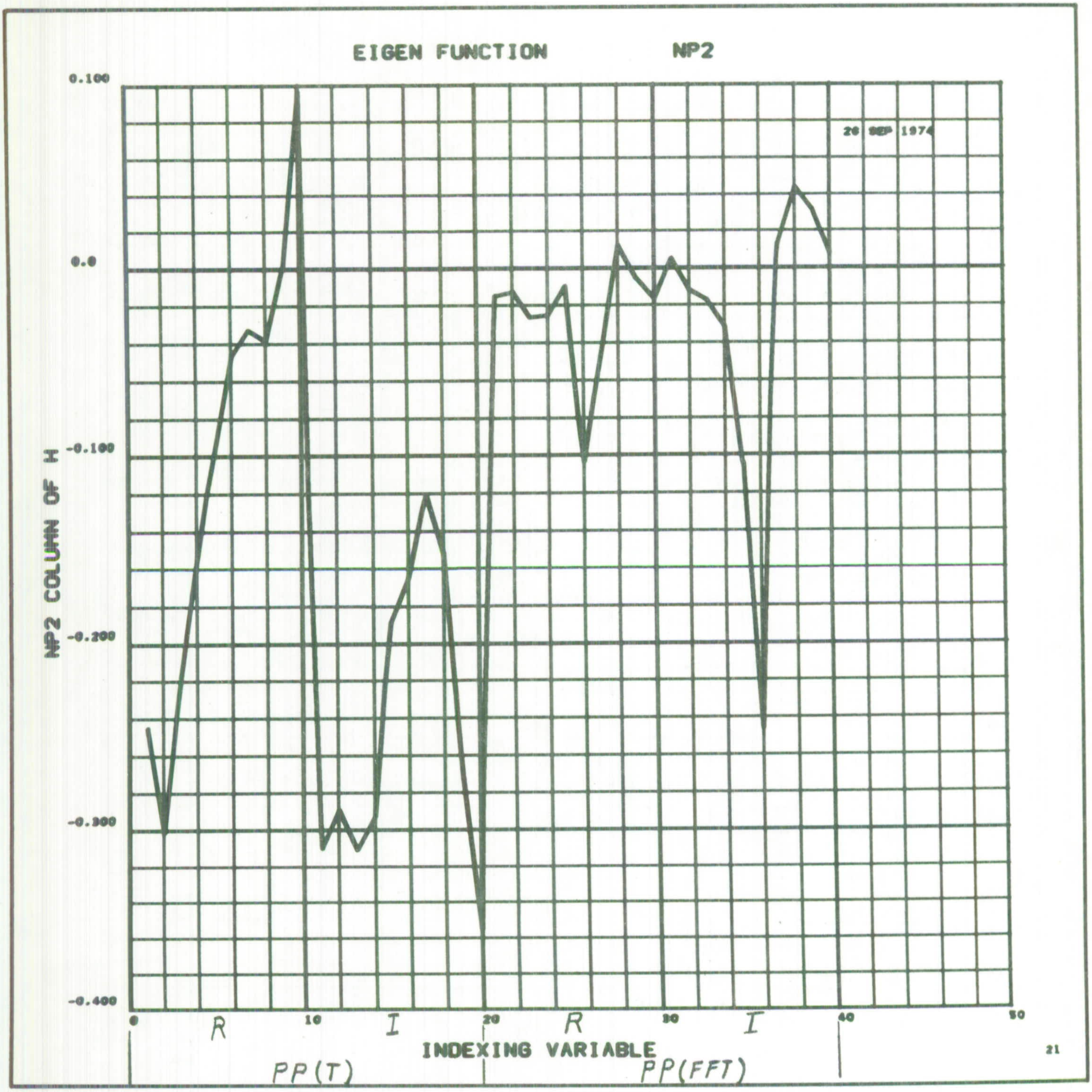
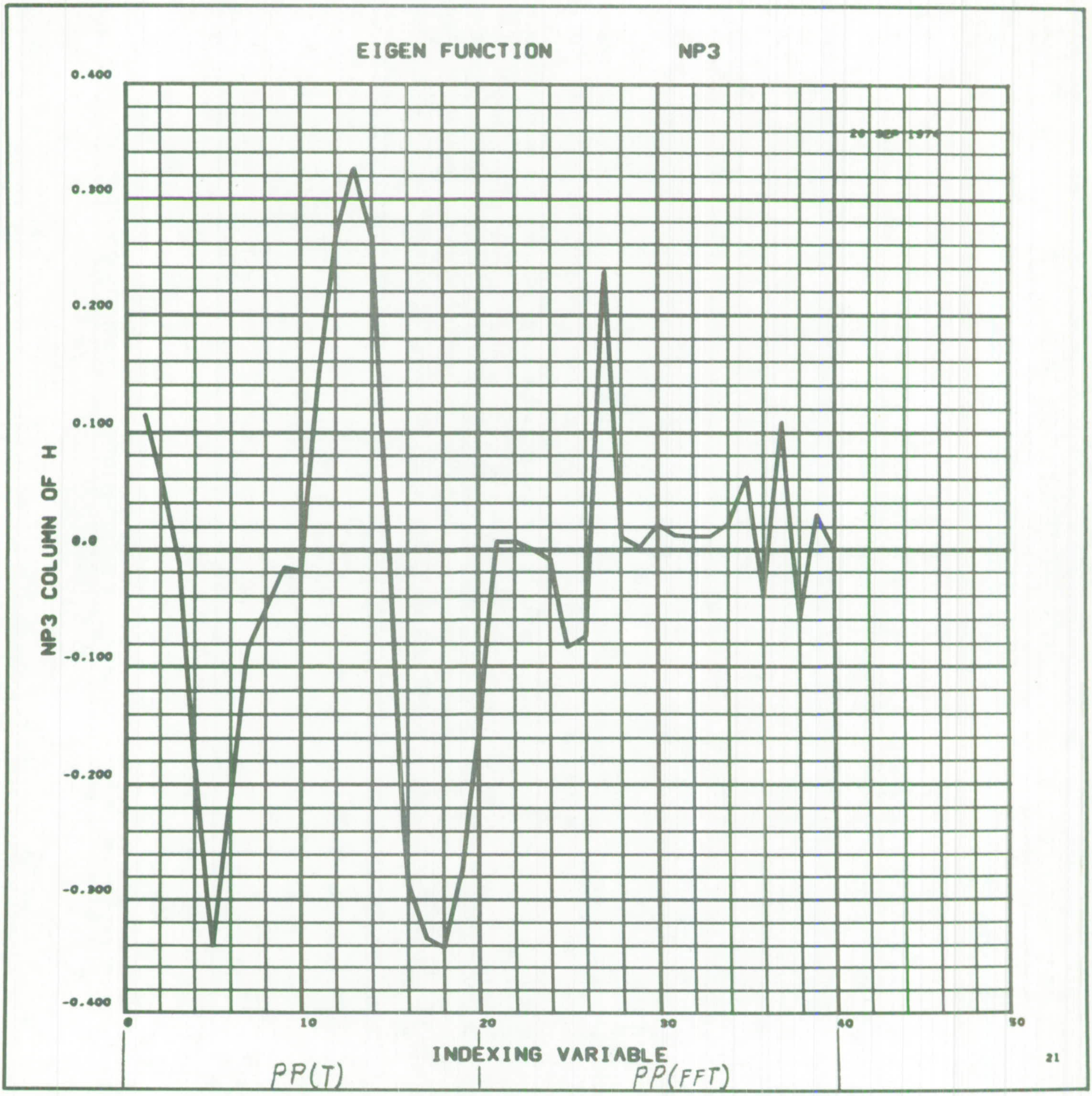


FIGURE 3.38 - THIRD OPTIMAL FUNCTION PRINCIPAL POLARIZATION BASE



as was contained in the fourth optimal function of the reference base. Examination of the higher order optimal functions associated with the principal polarization base indicate that there is no longer a 1 to 1 correspondence between this base and the reference base.

The scatter plot of the coefficients of the first two terms of the optimum generalized Fourier series expansion of the data histories in terms of the principal polarization base is presented in Figure 3.39. As would be expected from the discussion of the optimal functions, this figure is just a mirror image of the corresponding projection presented for the reference base presented in Figure 3.7. Examination of the higher order scatter plots show that for all the scatter plots the 101 targets occupy a single point at approximately the centroid of the scatter plot.

Thus, first three terms in the principal polarization base show great similarity to the corresponding terms in the reference base, however, significant differences develop after the third optimal function. This suggest that the gross features of the results obtained using only the principal polarization would be very similar to the gross features using both polarizations, but significant differences would occur after approximately the most correlated 30 to 40% of the information is used.

Time Base

The time base consists of a base constructed using only variables 1 through 20 and 41 through 60 from all of the data histories used in the reference base. The information energy or explained variation is a function of number of terms used in this base is presented in Figure 3.40. The most significant feature of this base is illustrated by the great similarity in distribution of information energy for the time domain only base and the reference base which was presented in Figure 3.4. In fact, comparison of Figures 3.4 and 3.40 shows no difference in the distribution of information energy. This suggests that the frequency domain is not adding any information over that which is available from the time domain. This is confirmed by comparing the optimal functions associated with the time base and the reference base. Comparison of Figure 3.41 with Figure 3.5 shows that the first optimal function of the time base is exactly the same as the corresponding points, that is

FIGURE 3.39 - SCATTER PLOT OF COEFFICIENTS OF FIRST VERSUS THE SECOND TERMS IN OPTIMAL REPRESENTATION-PRINCIPAL POLARIZATION BASE

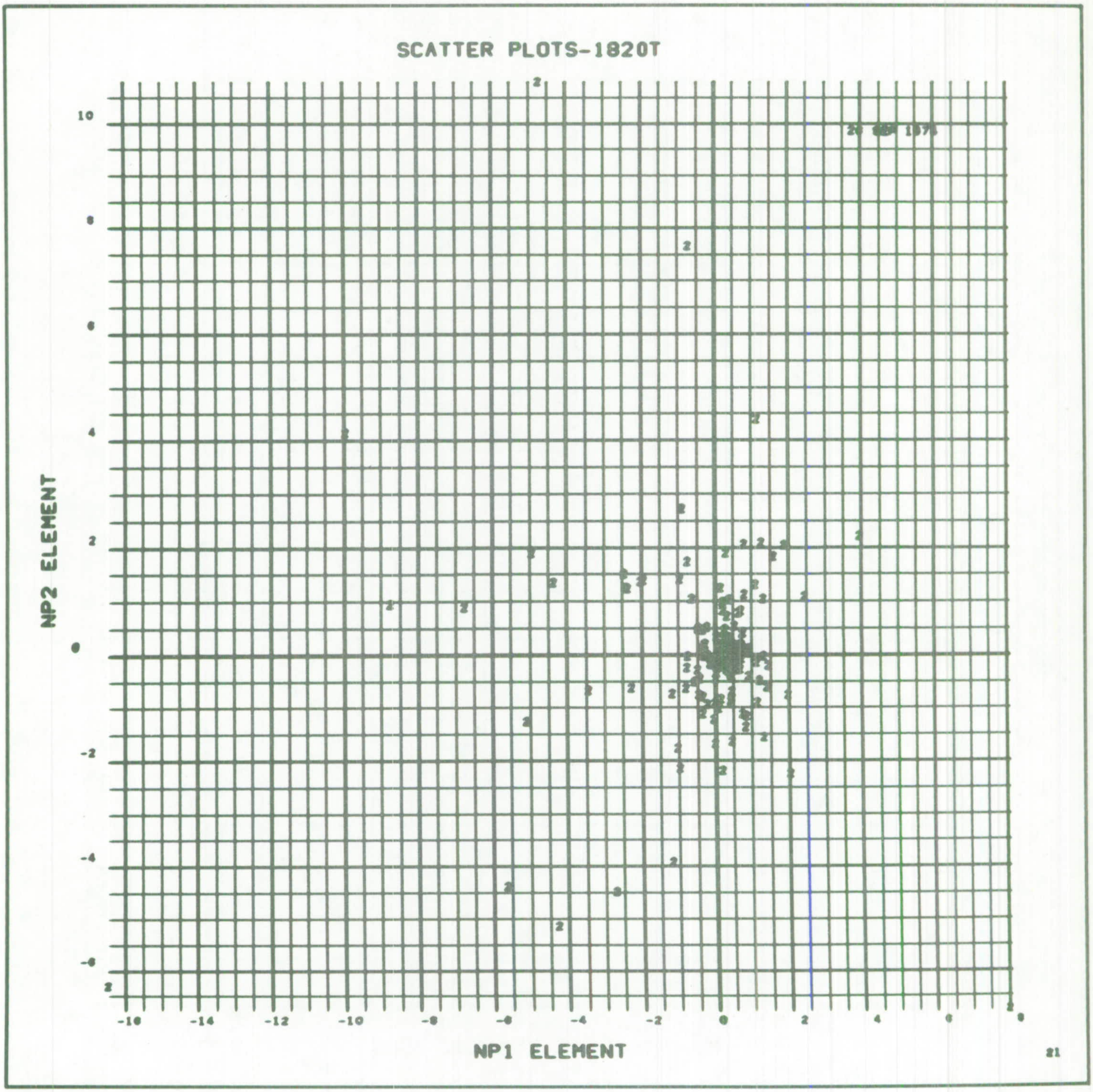


FIGURE 3.40- EFFECT OF DIMENSIONS ON AVAILABLE INFORMATION-
TIME BASE

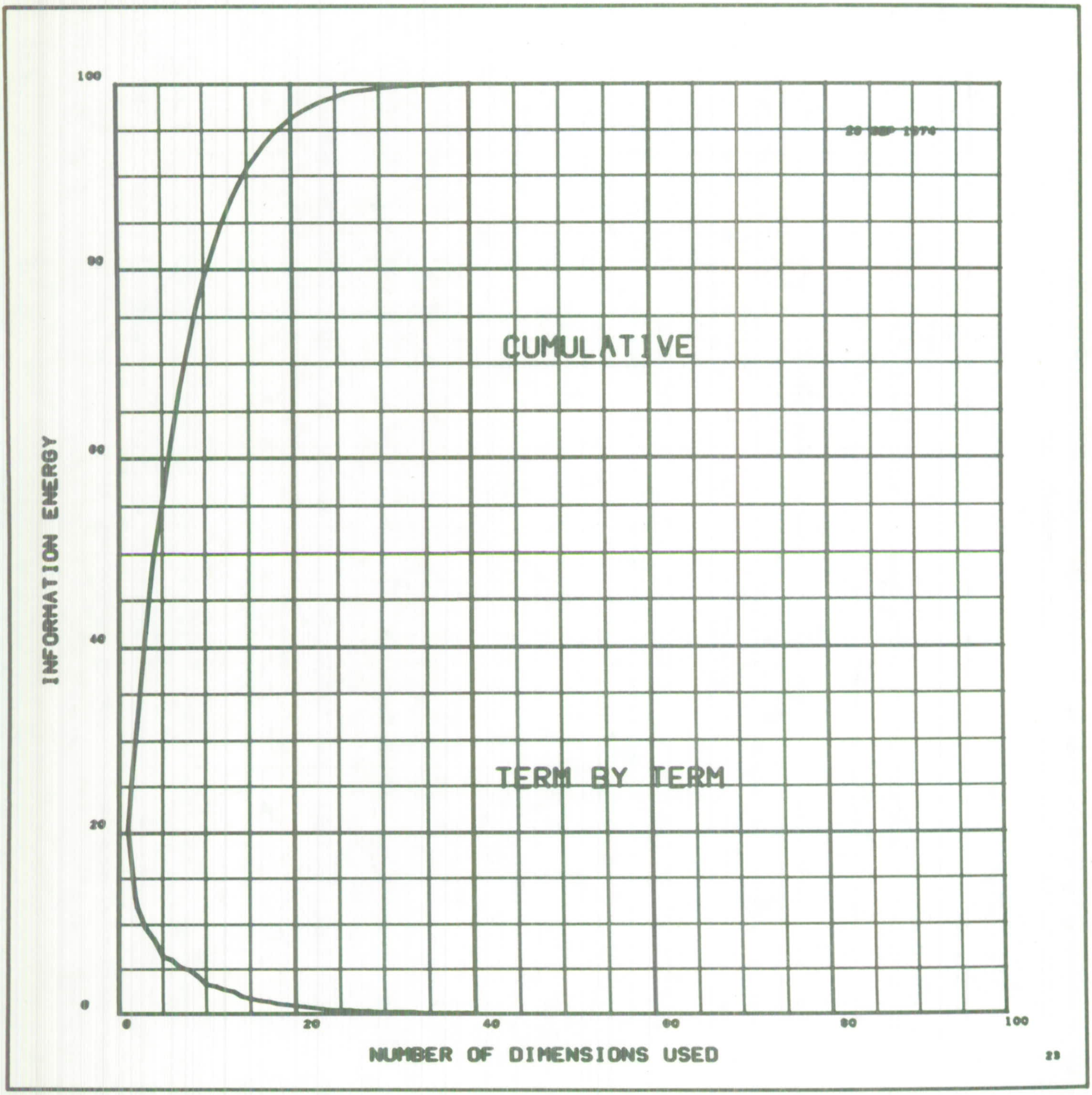
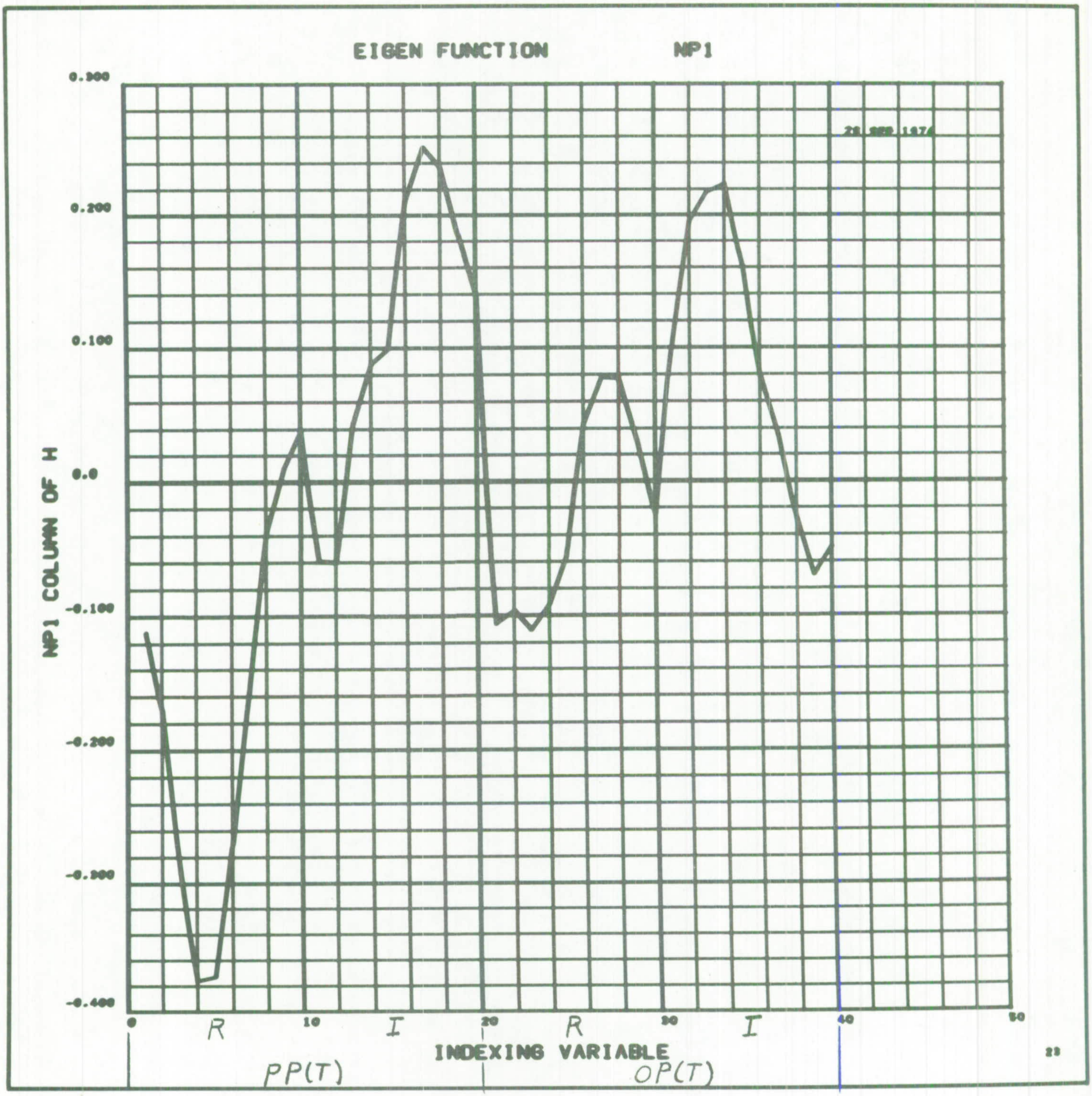


FIGURE 3.41 - FIRST OPTIMAL FUNCTION TIME BASE



1 through 20 and 41 through 60 of the reference base. The same statement is true for the second optimal function shown in Figure 3.42 when compared with the second optimal function of the reference based shown in Figure 3.6. Clearly, one would expect the scatter plots to be esentially identical and comparison of the scatter plot presented in Figure 3.43 with that presented in 3.7 for the reference base shows this to be the case. Detailed comparison of the higher order optimal functions shows that they are essentially identical through the maximum dimensionality used in the present analysis. To illustrate this, the fifteenth optimal function for the time domain base is presented in Figure 3.44. Comparison of this fifteenth optimal function with that presented in Figure 3.20 for the reference base shows that the fifteenth optimal function for the time base is identical to the regions covered by indexing variables 1 through 20 and 41 through 60 of the reference base.

This comparison of the reference and time only base verifies that the information contained in the real and imaginary components of the frequency plane is redundant with the information in the corresponding components of the time domain. This conclusion also follows from the linearity of the Fourier transform when one deals with the real and imaginary components. The remarkable result is not the similarity of the time portions of the optimal functions, but that the fact that the optimal functions for the reference base tended to show less variation in the frequency variables then in the time variables. It is especially surprising that this occurred in the dominant optimal functions.

It should also be emphasized that the conclusion regarding the identity of information content is only valid when one deals with the real and imaginary components of the time and frequency domain signatures. If one transforms to the amplitude and phase of the signatures the process

FIGURE 3.42 - SECOND OPTIMAL FUNCTION TIME BASE

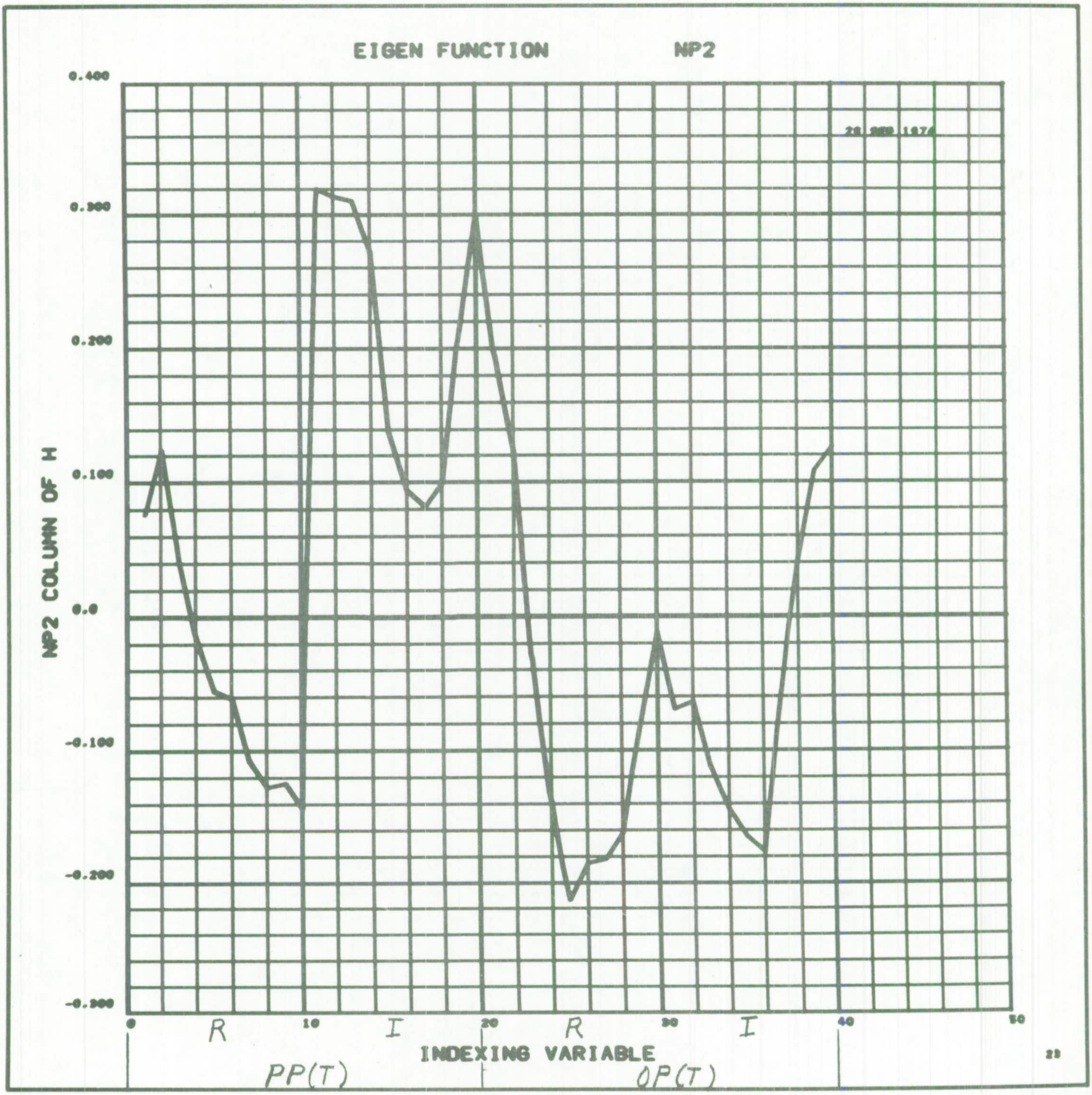


FIGURE 3.43 - SCATTER PLOT OF COEFFICIENTS OF FIRST VERSUS SECOND TERMS IN OPTIMAL REPRESENTATION-TIME BASE

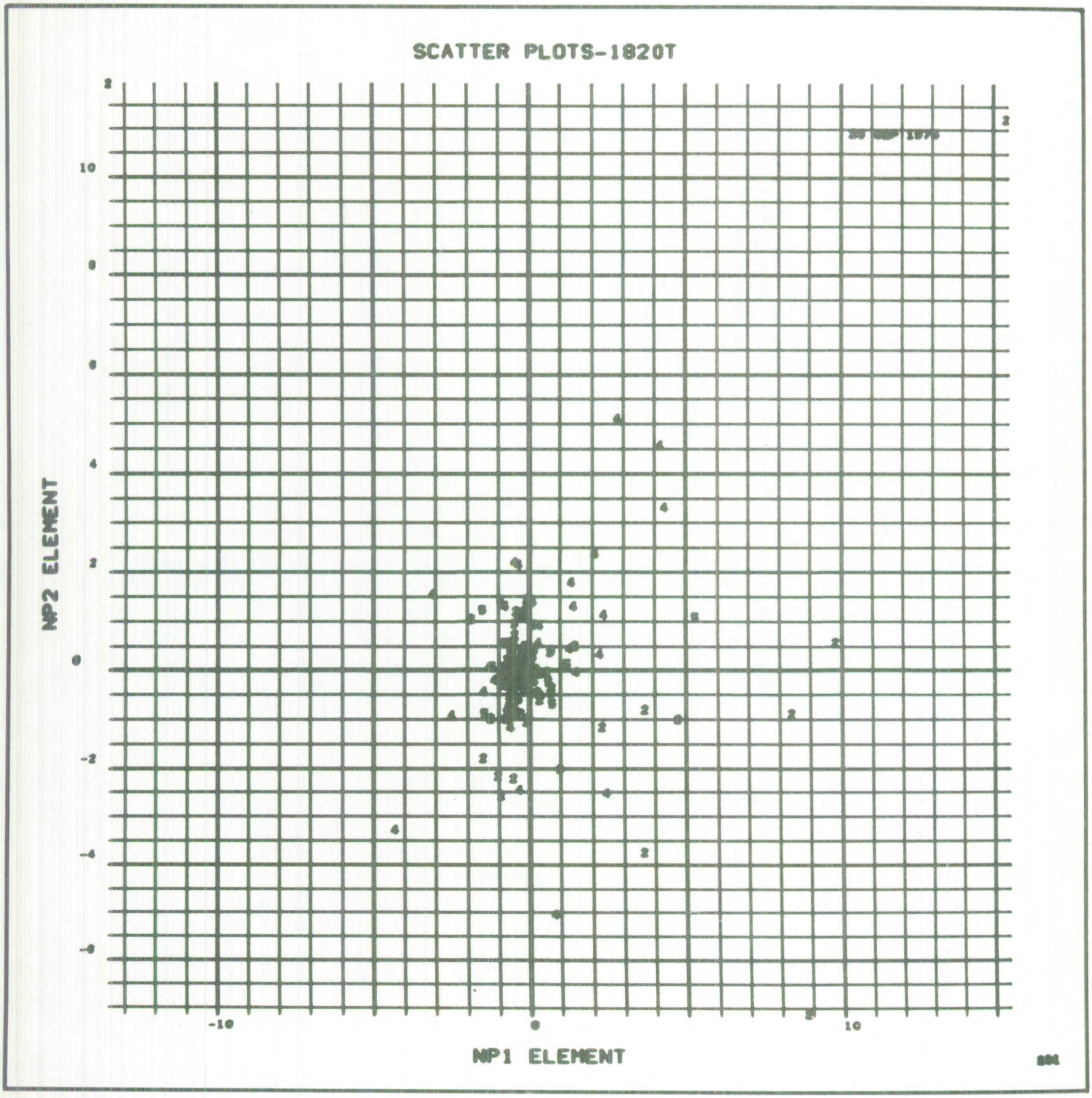
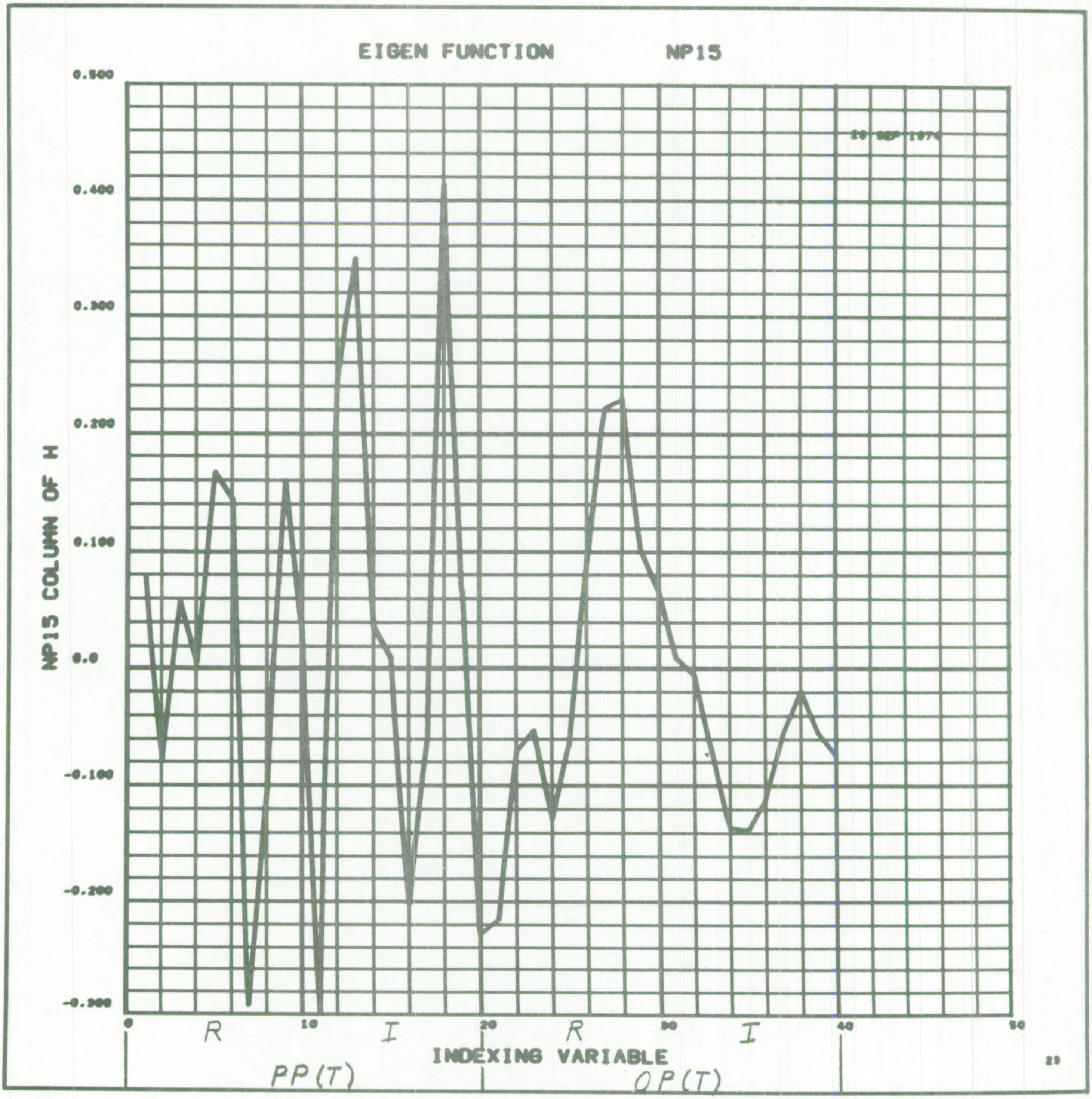


FIGURE 3.44 - FIFTEENTH OPTIMAL FUNCTION TIME BASE



becomes non-linear and the information content may be quite different. This was seen in the Task 1 studies where the amplitude in the time domain was found to be considerably less useful then the amplitude in the frequency domain.

Principal Polarization Time Base

The above comparison of the time domain and the frequency domain shows that one should construct the principal polarization base in the time domain alone. This base should contain the same information as the base constructed using the principal polarization from both the time and frequency domain. This would simplify the transformations to and from the optimal space and reduce the complexity of the data gathering system since only the principal polarized data must be collected and the Fourier transform would not have to be taken.

This base was constructed and designated the principal polarization-time base. Consistent with the above results, the information energy curve was identical to that presented in Figure 3.35. The optimal functions were identical to the first 20 points in the optimal functions for the principal polarization time base. Because of this identity, these figures are not presented but the reader is referred to the corresponding figures in the principal polarization base when interpreting results obtained using the principal polarization time base.

3.5 Bases Using Square Pre-Processing to Create Differences in Class Means

The configuration of the data points as indicated by the examination of the scatter plots suggested that if one took the square of the distance from the centroid of the 101 class to each target the resulting distribution of data points should be far more susceptible to classification by linear discriminants which use differences in the means of the classes. Since programs were not available to accomplish this in the optimal space and funding restrictions precluded modifying the programs to accomplish this and alternate approach of squaring the components in the original data space and then finding the optimal representation was tried. Two bases were made utilizing essentially the square of the data vector. The first subtracted the mean of Class 1 from the data prior to squaring the components, the other base was made by simply squaring the components in the data space prior to processing. The rationale for squaring

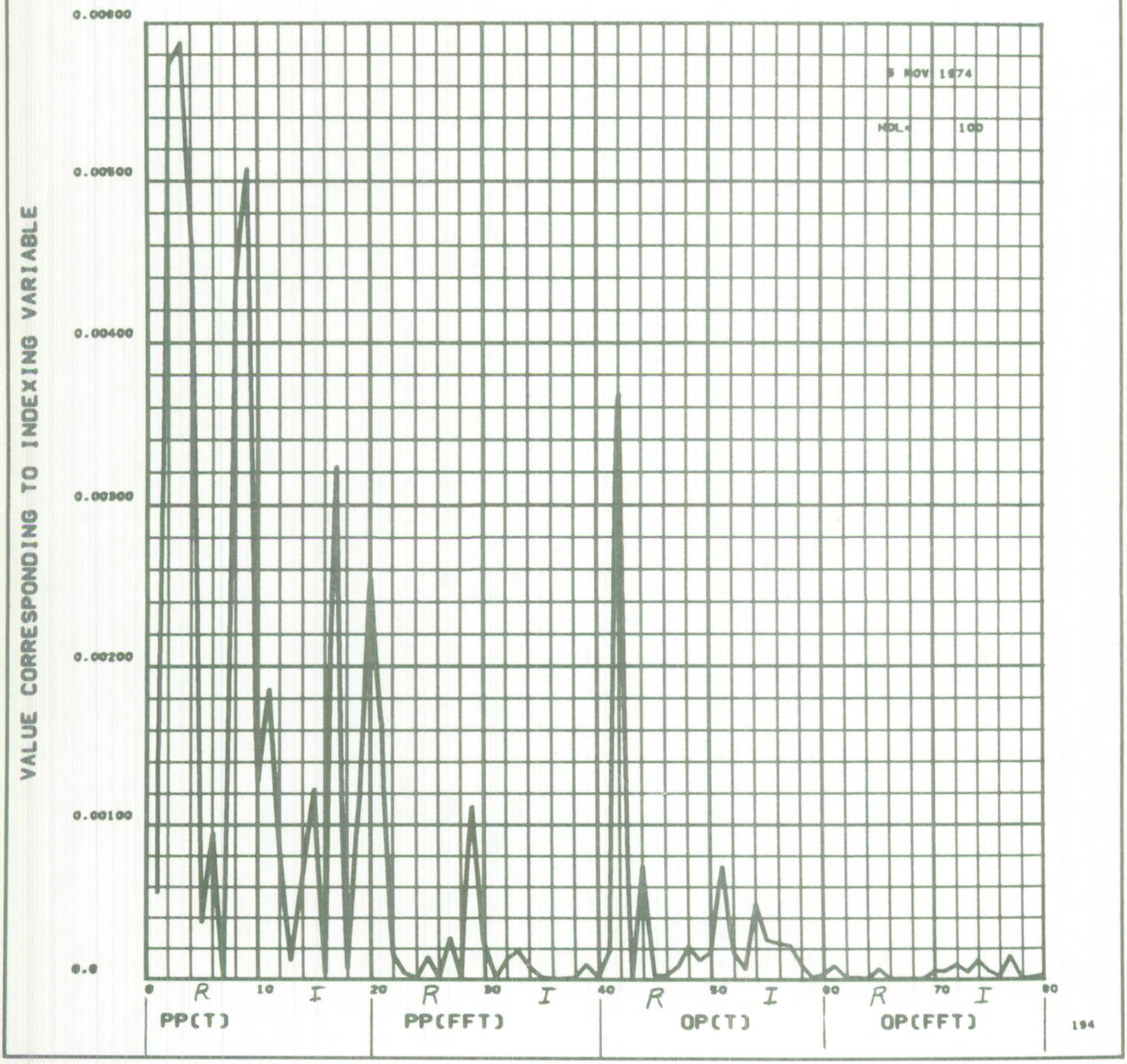
the data without subtracting the mean is based on the extremely small value of the mean for Class 1. The results indicated that both approaches produced essentially the same base.

Typical data histories for the same cases as shown in Figures 3.1 and 3.2 are shown in Figures 3.45 and 3.46 using the square of each of the components. The average of all 260 data histories used to make the square base is presented in Figure 3.47. The average of the data histories prepared with or without prior subtraction of the mean was indistinguishable. The same was true of the distribution of information energy as a function of number of terms used. This distribution is given in Figure 3.48 and shows that the square variables are significantly easier to represent than the reference base. For example, the first term in Figure 3.48 contains 43% of the information in the entire data set. The 15 term analysis using the square base utilizes 96% of the total information as compared to 91% for the base using the variables without the square pre-processing.

The first and second optimal functions using the square variables are presented in Figures 3.49 and 3.50. The scatter plot of the coefficient of these functions for each of the data histories used is presented in Figure 3.51. Note, that in this figure we see the same characteristic that the 101 class again appears as a single point and only some of the 21X class points are not located at this point. However, comparing Figure 3.51 with the reference scatter plot presented in Figure 3.7 shows that the squaring processes has resulted in moving the point representing the 101 class from the centroid of the data to one side of the data. Thus, the means of the two classes are no longer identical and one might expect some of the linear discriminants which make use of the difference of means to be more effective on this data.

The insignificance of the prior subtraction of the mean of the 101 class was illustrated by the near identity of all of the optimal functions from the first through the fifteenth. The most dissimilar functions can be expected to be the higher numbered functions where the noise plays a more significant roll. Thus, the best illustration of this similarity is to compare the fifteenth optimal functions which is done in Figures 3.52 and 3.53. These functions are mirror images of one another but other than the mirror imagery

FIG - 3.45 SAMPLE 10 PULSE SQUARE HISTORY FOR TARGET 101



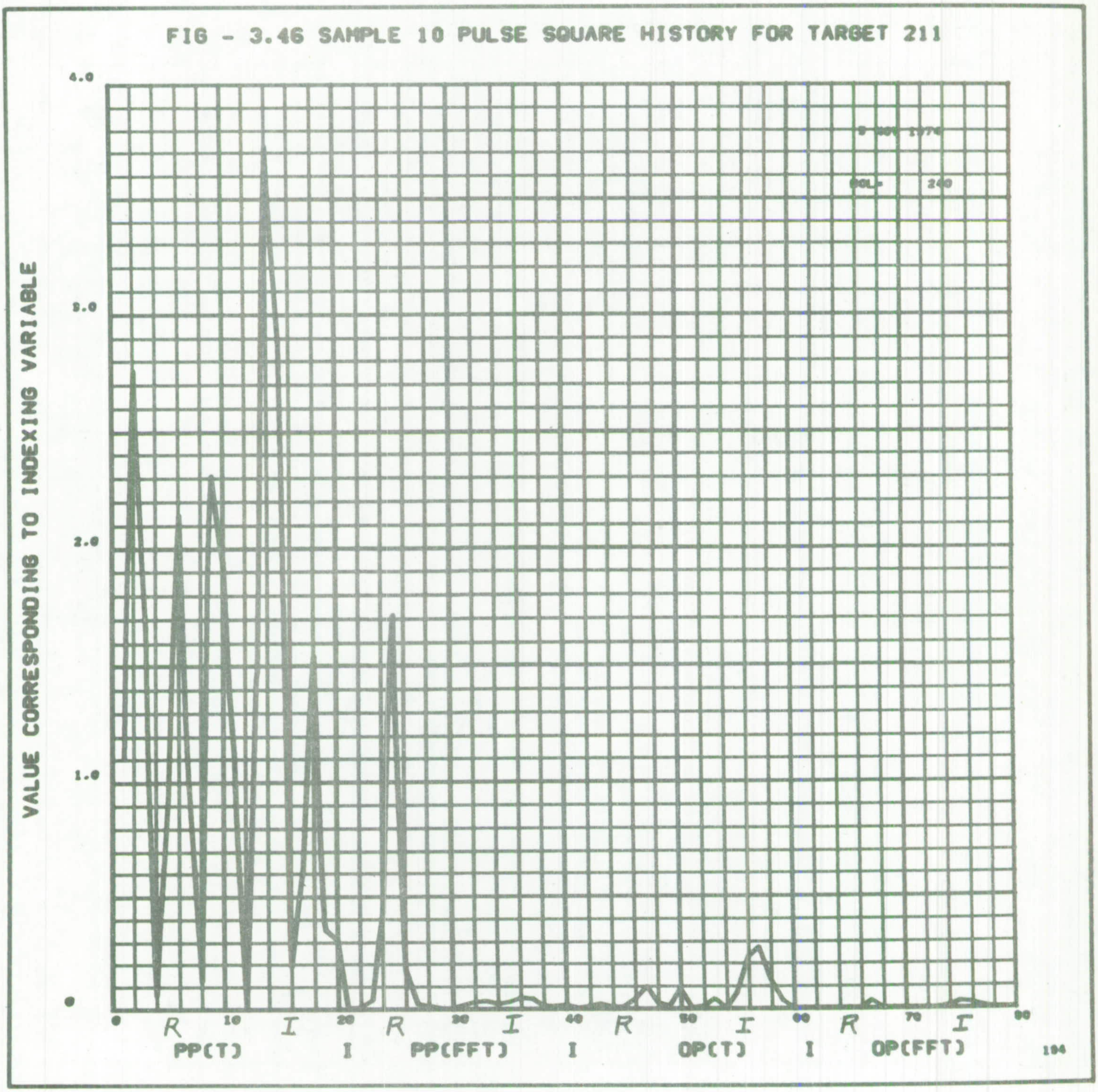


FIG - 3.46 SAMPLE 10 PULSE SQUARE HISTORY FOR TARGET 211
VALUE CORRESPONDING TO INDEXING VARIABLE
4.0
3.0
2.0
1.0
0
10
20
30
40
50
60
70
80
R I R I R I R I
PP(T) I PP(FFT) I OP(T) I OP(FFT)
194

FIGURE 3.47 - AVERAGE OF ALL 260 CASES USED TO CONSTRUCT SQUARE BASE

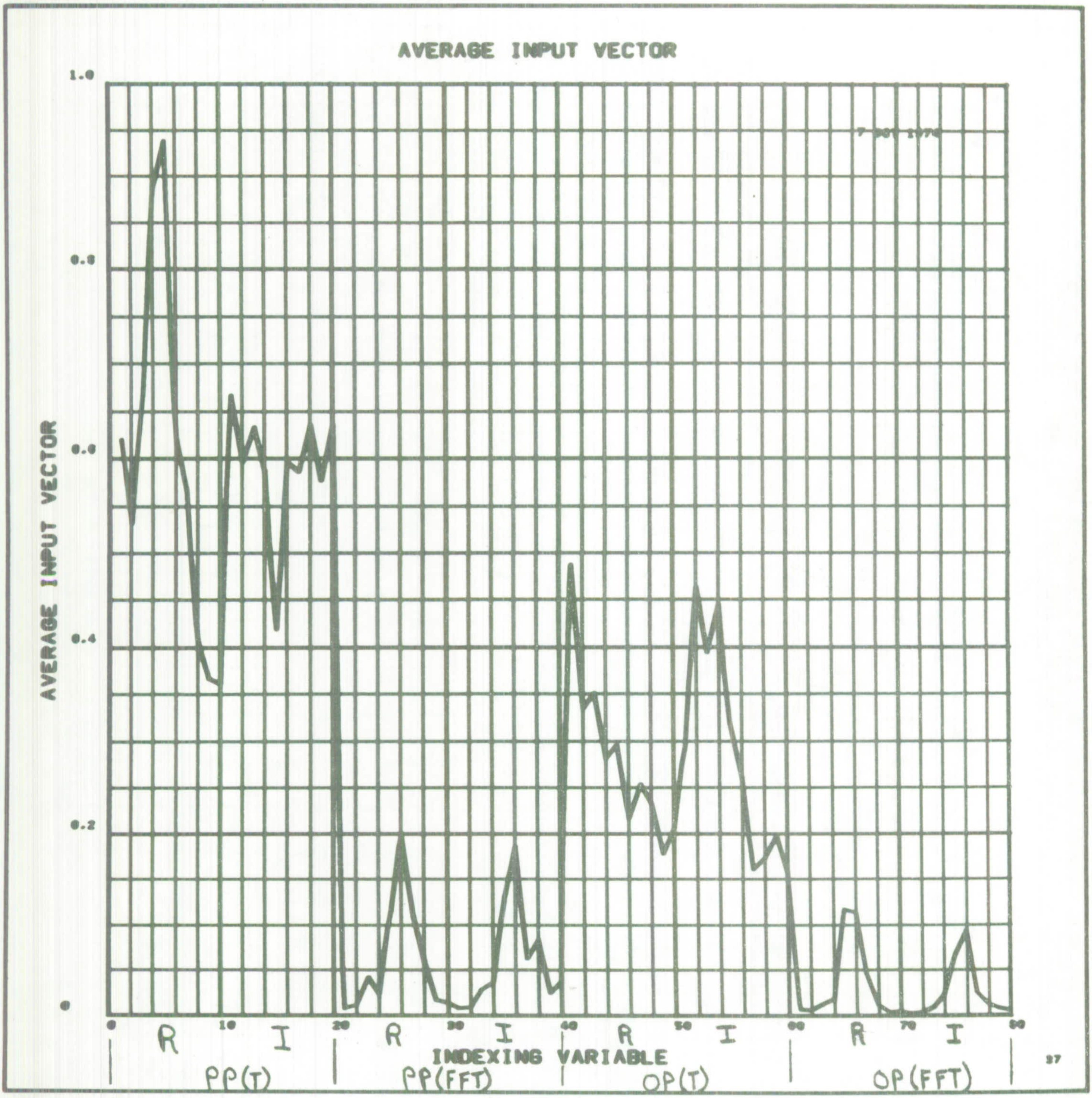


FIGURE 3.48 - EFFECT OF DIMENSIONS ON AVAILABLE INFORMATION-
SQUARE BASE

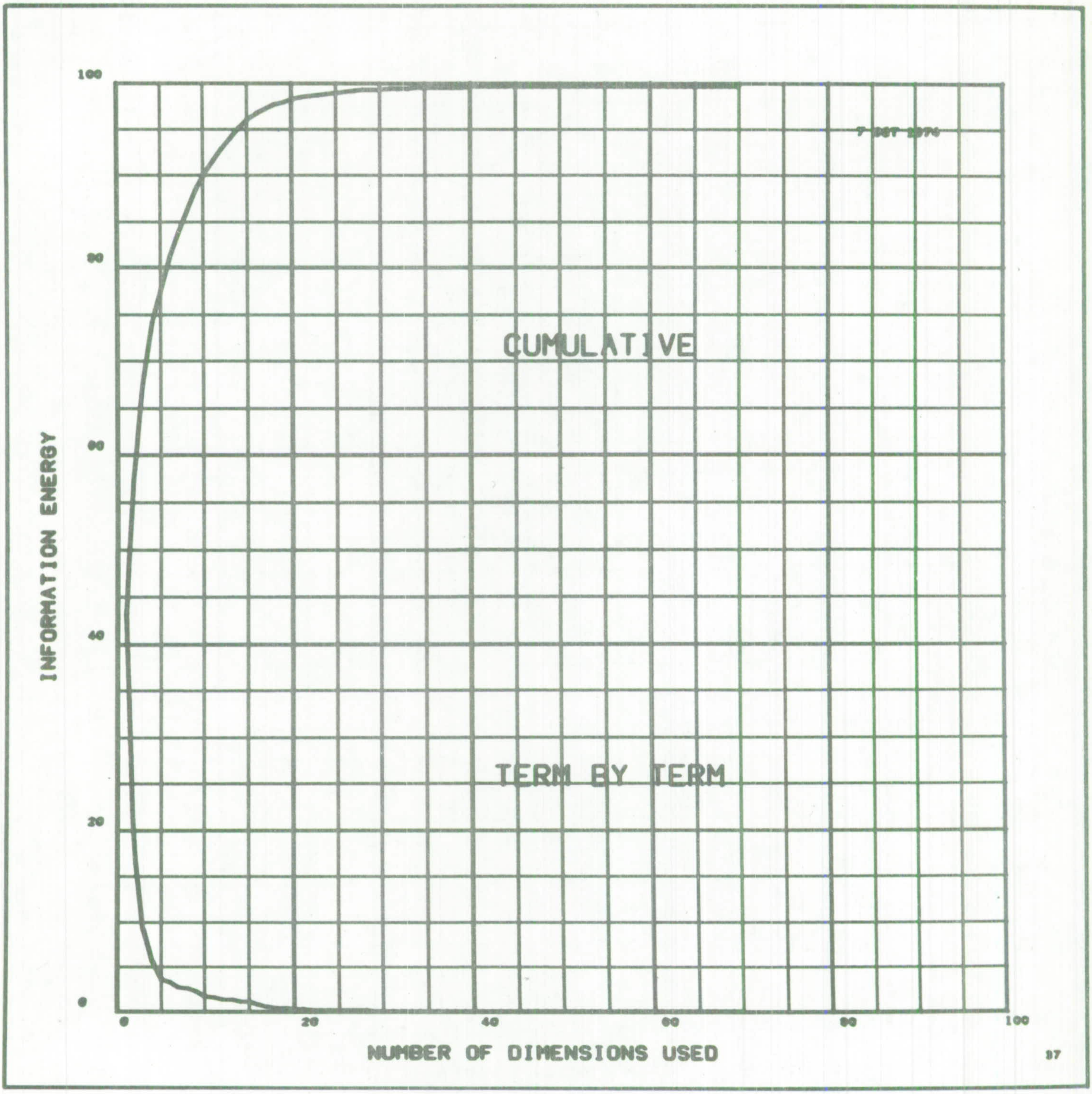


FIGURE 3.49 - FIRST OPTIMAL FUNCTION-SQUARE BASE

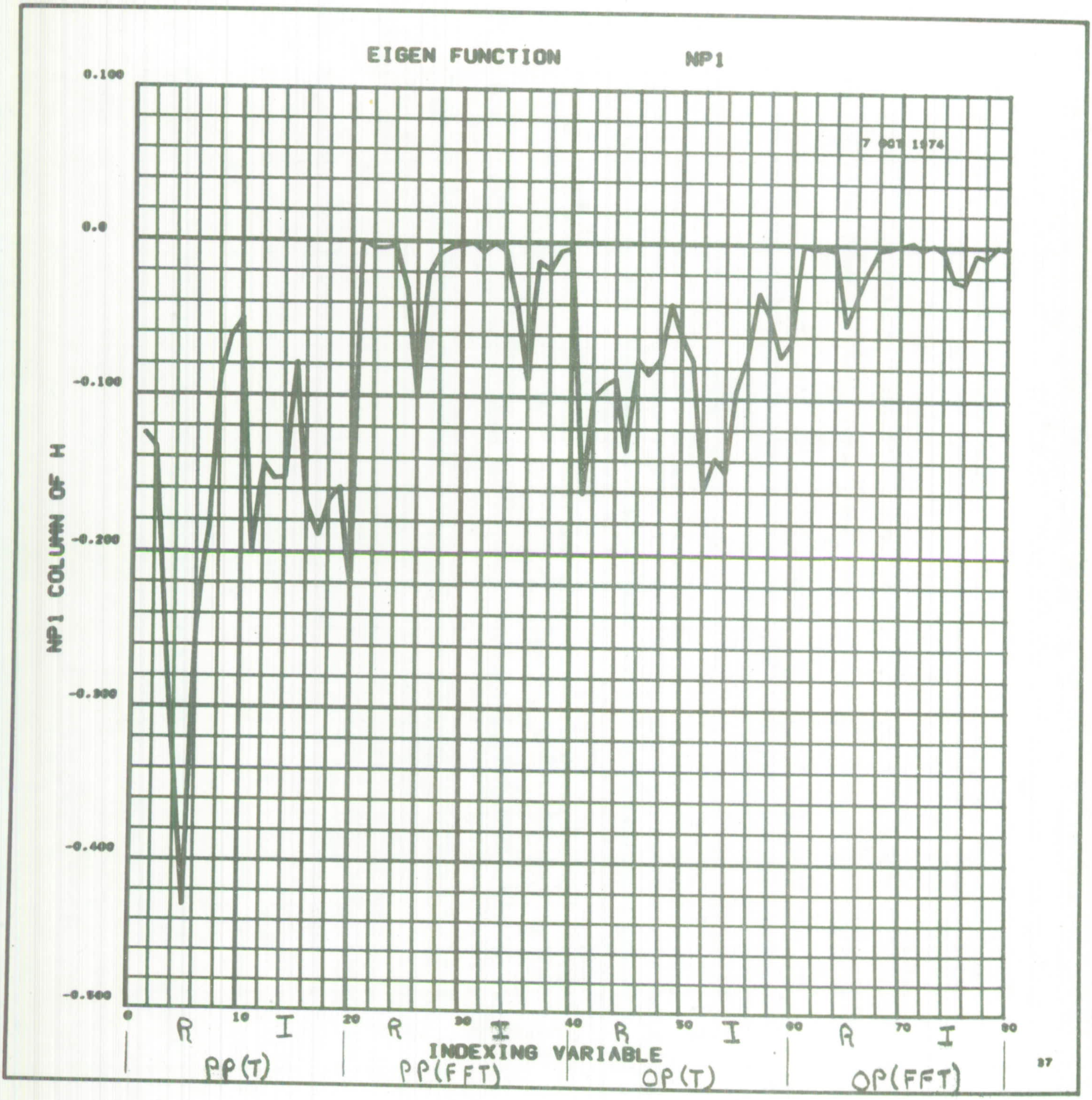


FIGURE 3.50 - SECOND OPTIMAL FUNCTION-SQUARE BASE

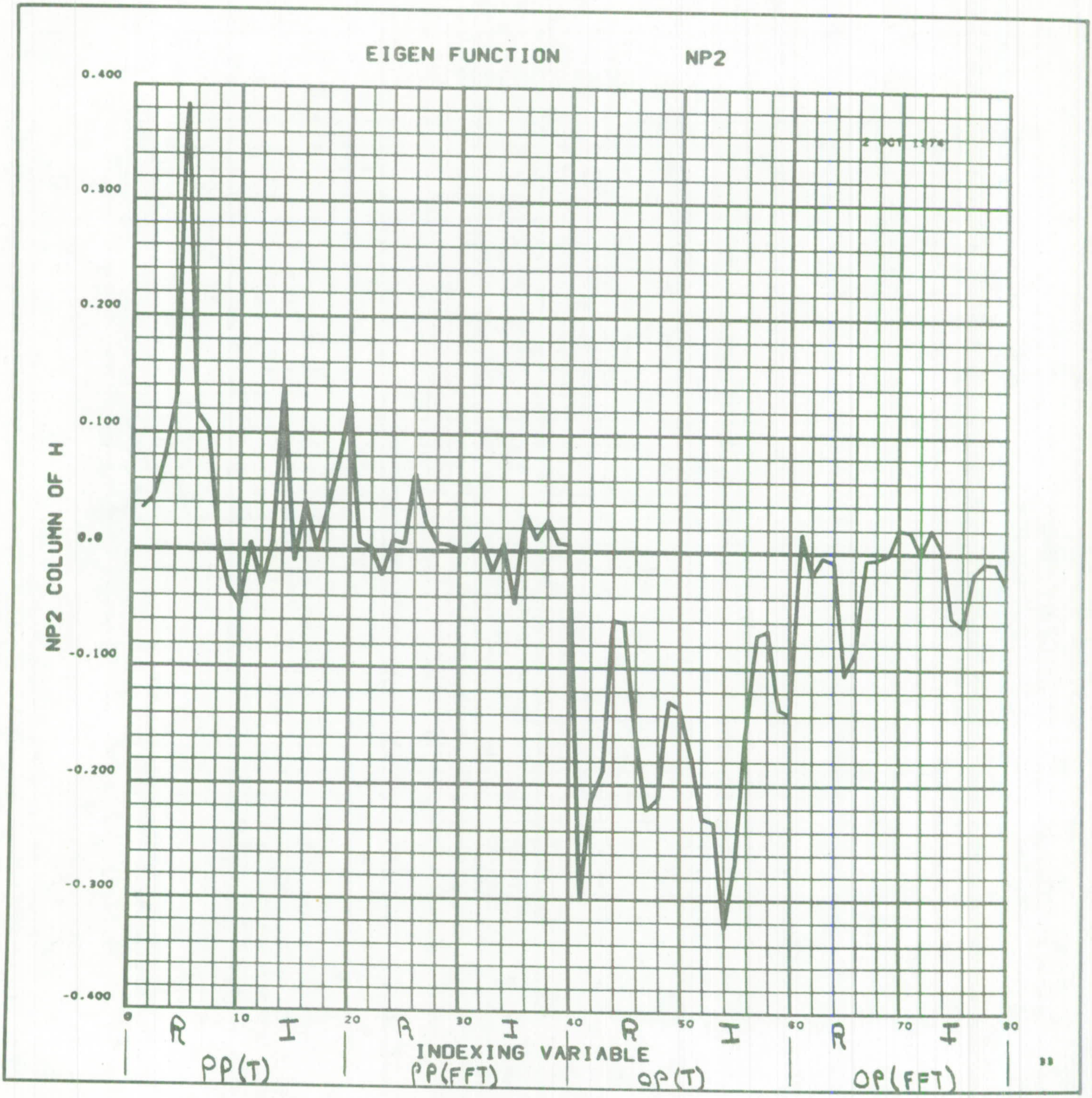


FIGURE 3.51 - SCATTER PLOT OF COEFFICIENTS OF FIRST VERSUS SECOND TERMS IN OPTIMAL REPRESENTATION-SQUARE BASE

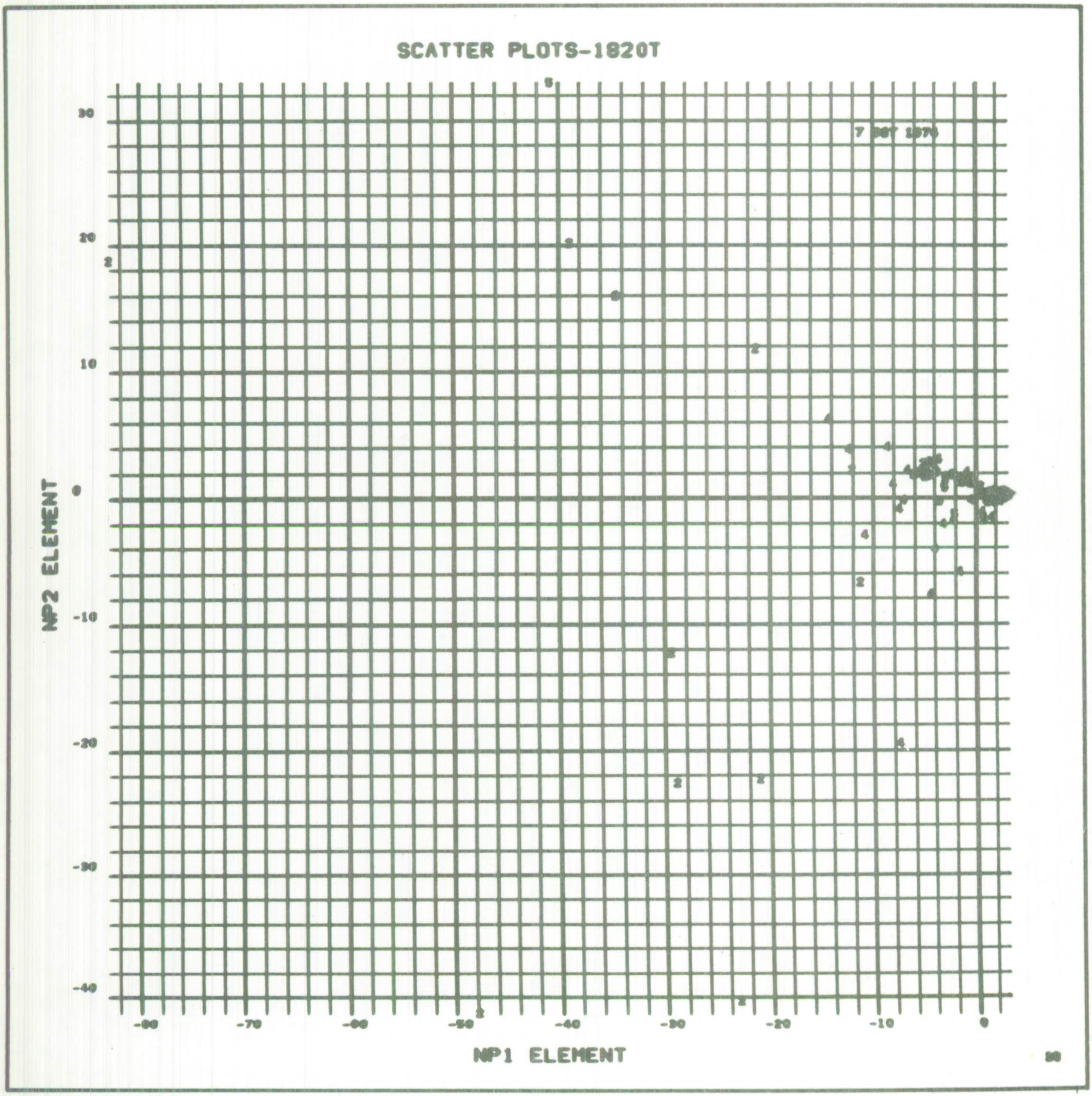


FIGURE 3.52 - FIFTEENTH OPTIMAL FUNCTION ZERO MEANS SQUARE BASE

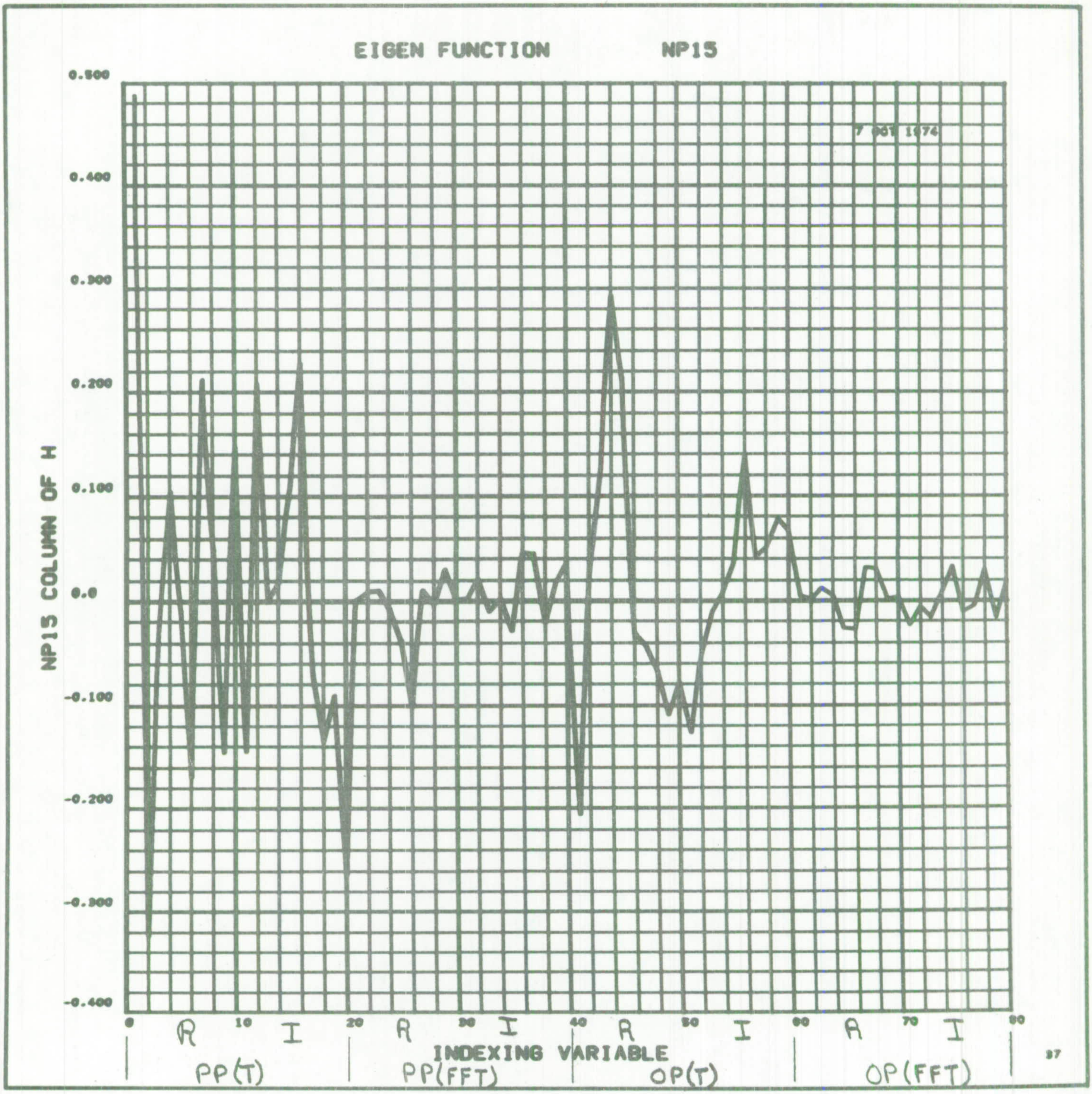
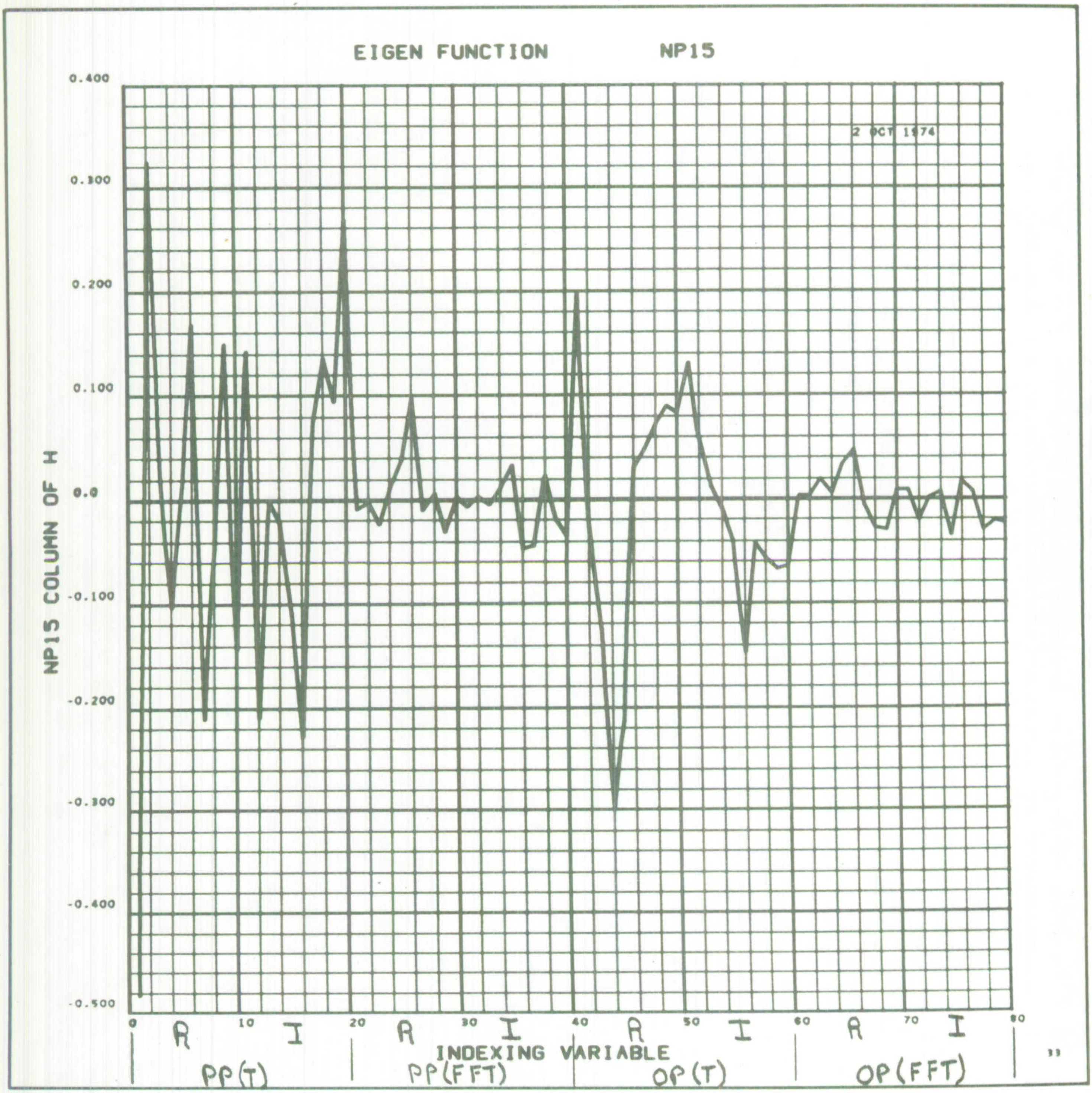


FIGURE 3.53 - FIFTEENTH OPTIMAL FUNCTION SQUARE BASE



the two functions are nearly identical. This verifies that the mean of the 101 class can be taken as zero without significantly effecting the results of the analysis.

4.0 ANALYSIS OF LINEAR CLASSIFICATION SCHEMES

In this report linear classification schemes are considered to be any scheme having a linear detection statistic for which the primary computation required may be visualized as the projection of the data vector on to another vector. This operation mathematically is the DOT product operation and thus simply requires the correlation of the two vectors. Because of this computational simplicity, linear classification schemes as defined here can be implemented with much simpler hardware or for a given computer capability many more cases can be examined than for non-linear schemes. In this section of the report, we shall review the various linear classification schemes which have been investigated for separating the Class 101 targets from the Class 21X targets.

4.1 Selection of Candidate Linear Classifiers

The selection of the linear classifiers can be aided by the analysis of the ADAPT representation which was presented in Section 3. The scatter plot projections for all of the learning data as well as 340 independent test cases are shown in Figure 4.1. In this figure, Numerals 1 and 2 represent the one hundred learning and one hundred test cases for the 101 class, respectively. The odd Numerals 3 through 9 represent the 40 learning cases for each of the four 21X classes used in this study. The even Numerals 4 through 8 and the equal signs represent the 60 test cases for each of the four 21X classes.

This figure shows that the learning and test data both exhibit the general characteristics of the scatter plots which were seen on the learning data plot presented in Section 3. The 101 Class points are so densely packed that they are not visible on this figure. Comparison of the odd and even numbers indicates that the outlier cases are equally divided between learning and test cases from similar classes. This general character is true in all of the higher order scatter plots. The expansion of the dense central region for the learning cases is presented in Figure 4.2. In this figure, the one hundred Class 101 learning cases are indicated by the Numeral 1. The sixteen Class 21X learning cases closest to the centroid of the 101

FIGURE 4.1 - SCATTER PLOT OF COEFFICIENTS OF FIRST VERSUS SECOND TERMS IN OPTIMAL REPRESENTATION ILLUSTRATING THE PROJECTION OF ALL 600 LEARNING AND TEST CASES ON THE REFERENCE BASE

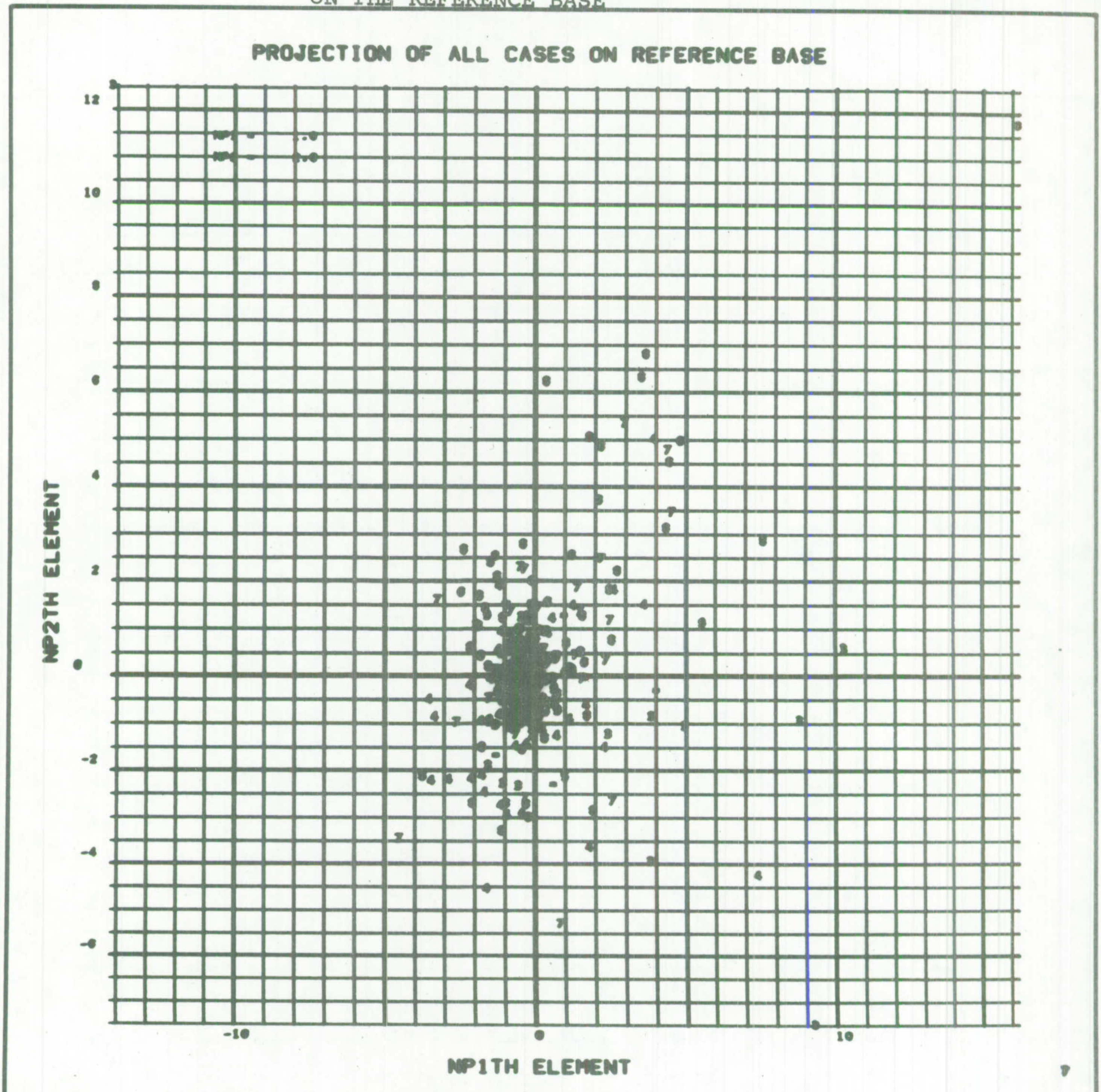
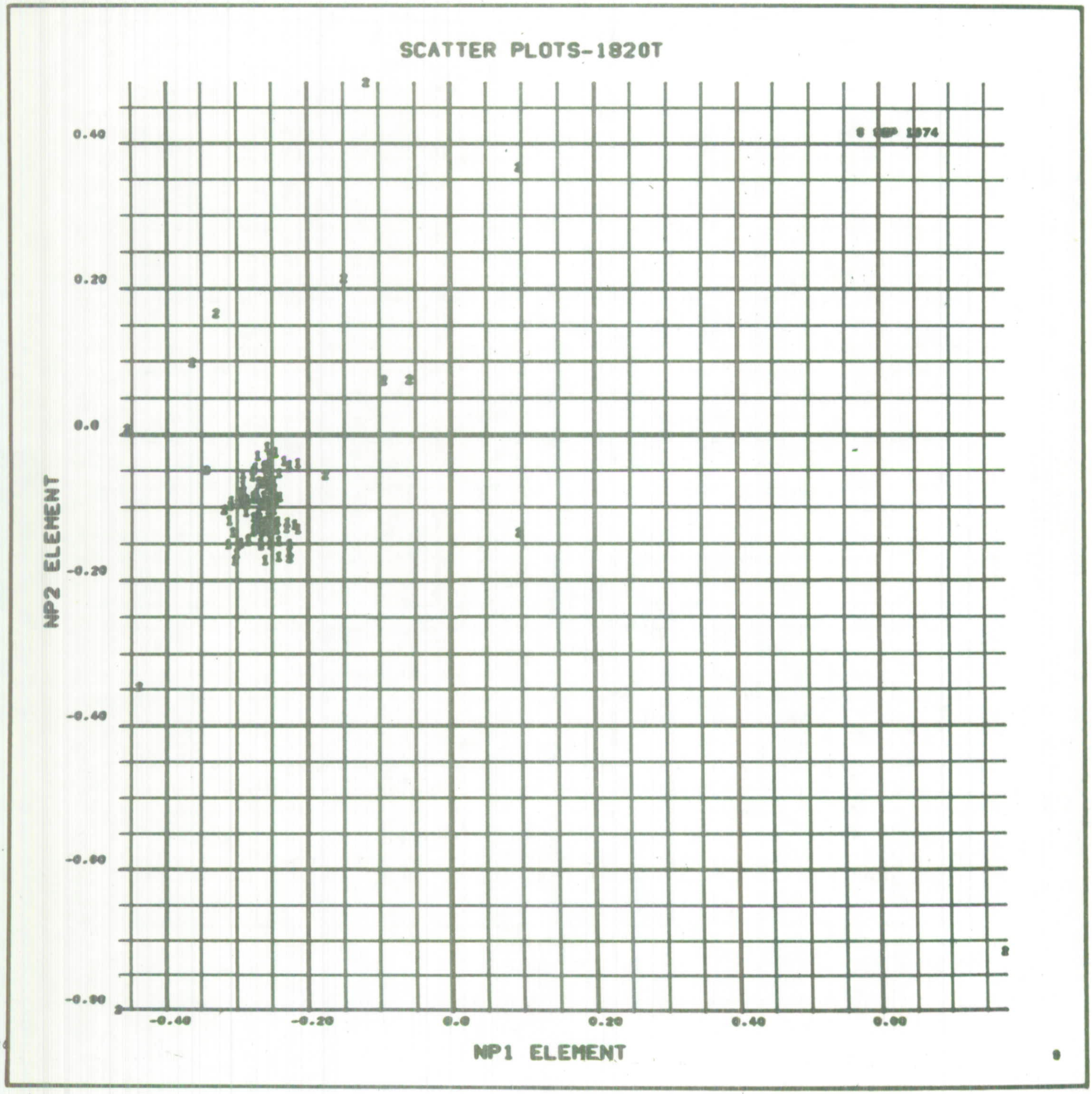


FIGURE 4.2 - SCATTER PLOT OF COEFFICIENTS OF FIRST VERSUS SECOND TERMS IN OPTIMAL REPRESENTATION SHOWING THE LOCATION OF LEARNING CASES NEAR THE 101 CLASS ON THE REFERENCE BASE



classes are indicated by the Numerals 2. This plot verifies the hypothesis that the 101 class is very densely packed in this space relative to the 21X Class. This characteristic is also true in the higher order scatter plots. An example of such a scatter plot, the plot of the tenth and eleventh optimal functions is presented in Figure 4.3.

These plots show that the major difference between the 101 and 21X Class is the variance while both classes have similar means. This result is very important in selecting the linear classifier[1] to be used to separate the classes. In most classification problems to which the ADAPT programs have been applied, the difference between the classes was to a large extent related to the characteristics of the mean of the classes. For these cases, the Fisher discriminant has proved both effective and easy to implement. However, when the mean values of two classes are similar and the variance provides the basis for discrimination, the Fisher discriminant has serious difficulties. This is illustrated by performance of the Fisher discriminant on the learning data which is shown on Figures 4.4A through C. Figures 4.4A and B present the projection of the Class 21X learning data onto the Fisher classification direction. Figure 4.4C presents the projection of the Class 101 learning data onto the Fisher direction. Examination of this figure shows that even though the means of the two classes have been separated by the Fisher discriminant, the great variation in the 21X Class has resulted in false alarm rates approaching 50%.

The solution to this problem is illustrated by the diagram presented in Figure 4.5. The class having very small variation represented by the solid curve in the center of Figure 4.5 may be entirely enclosed within two thresholds. Although the class having a large variance will have a few cases falling between these thresholds, the majority of its members will fall outside of this limited region. The bar chart presented in Figure 4.4A shows that even with the Fisher discriminant considerably better performance would be obtained if one used two threshold values. If one identified the 101 Class as all of those cases whose projection on the Fisher direction fell between values of 0.1 and 0.2 and all other values regardless whether they were above or below this value were considered to be members of Class 21X, the performance would be considerably improved. The emphasis in selection of the Fisher direction is on both the mean and variation of

---

(1) Note that definition of linear classifier used in this report includes any classifier for which the detection statistic is obtained in a linear manner.

FIGURE 4.3 - SCATTER PLOT OF COEFFICIENTS OF TENTH VERSUS ELEVENTH TERMS IN OPTIMAL REPRESENTATION-SHOWING THE LOCATION OF LEARNING CASES NEAR THE CENTROID OF THE 101 CLASS

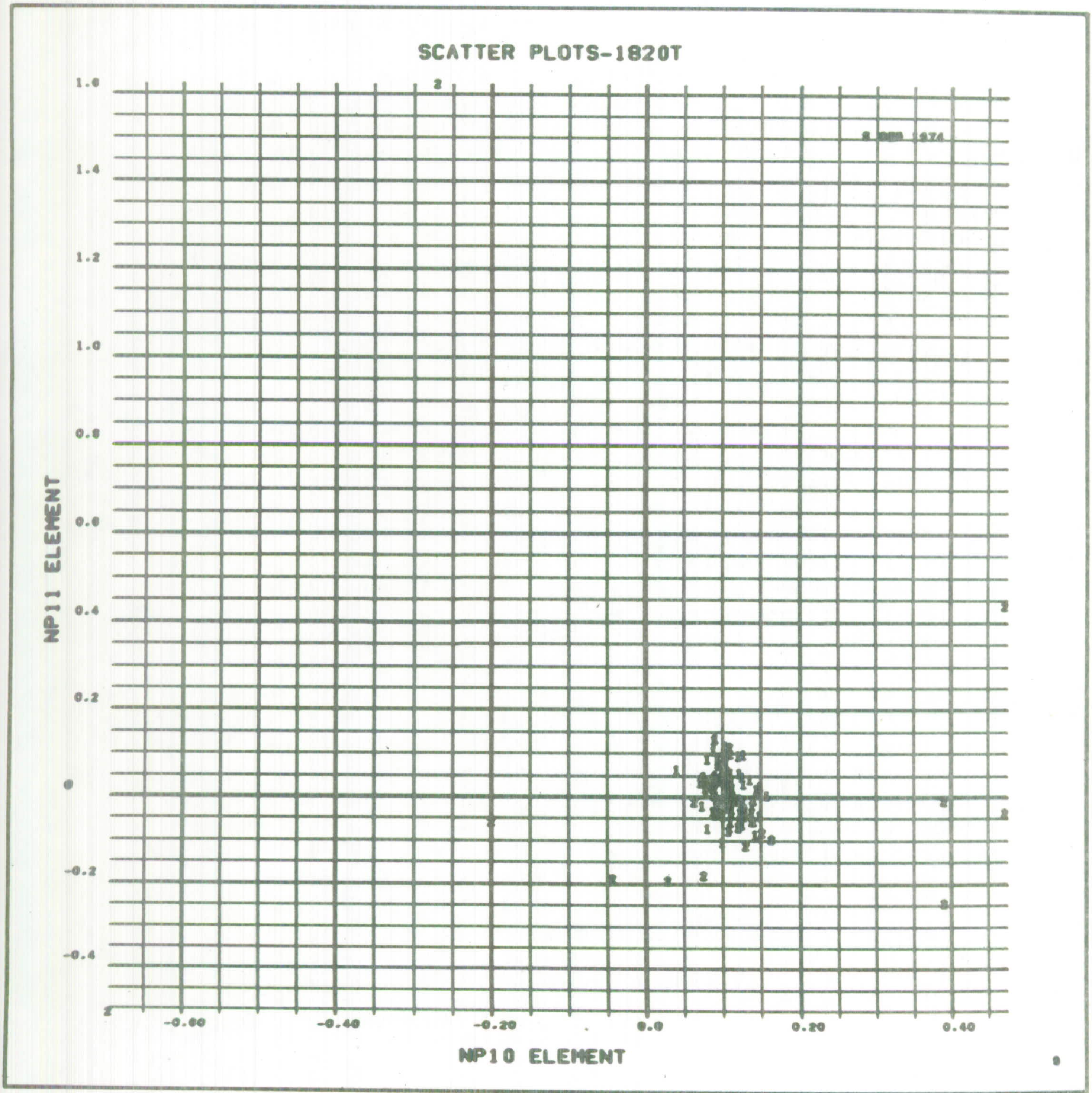


FIGURE 4.4A- BAR CHART SHOWING MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE FISHER CLASSIFICATION DIRECTION USING THE REFERENCE BASE

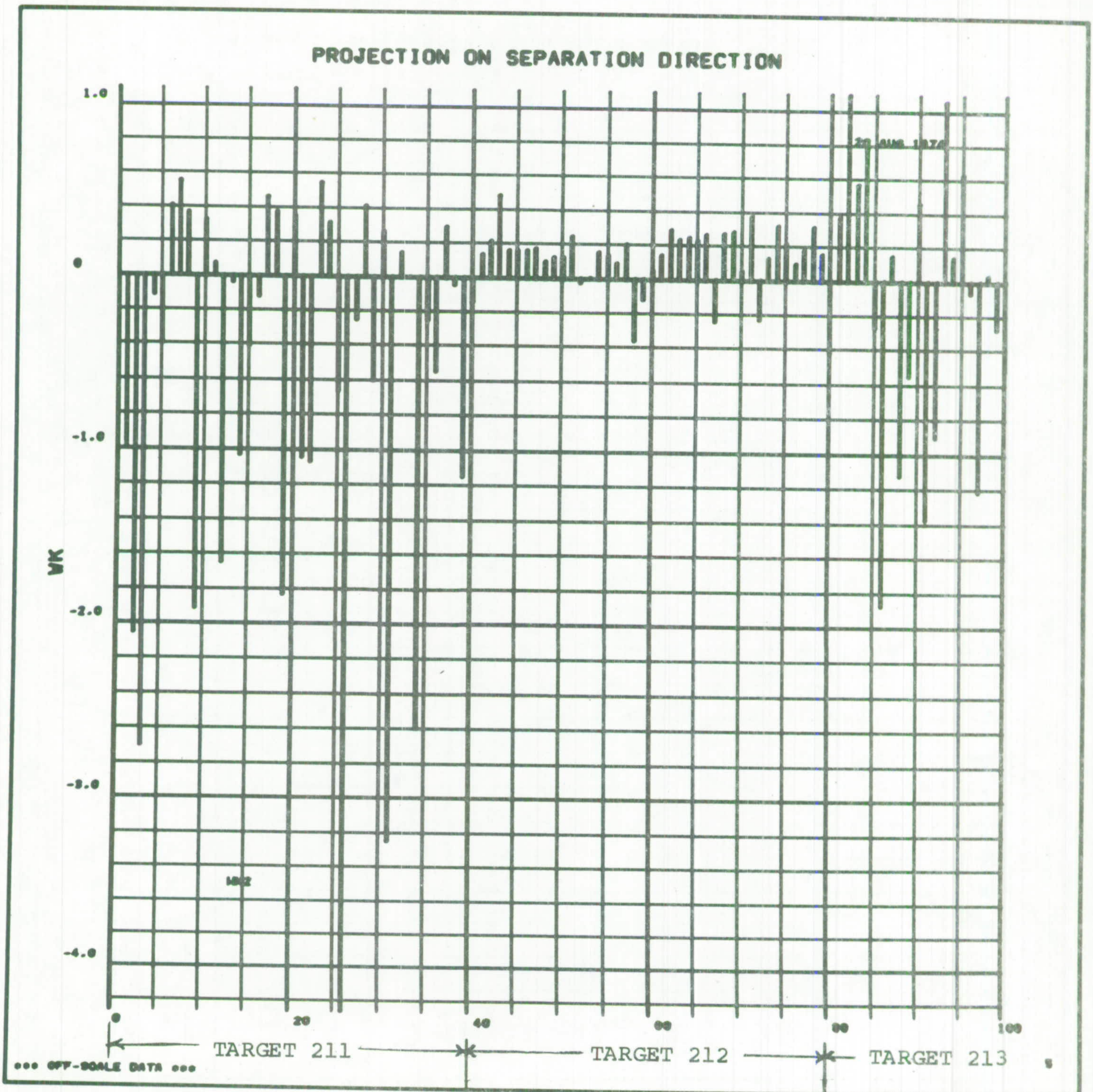
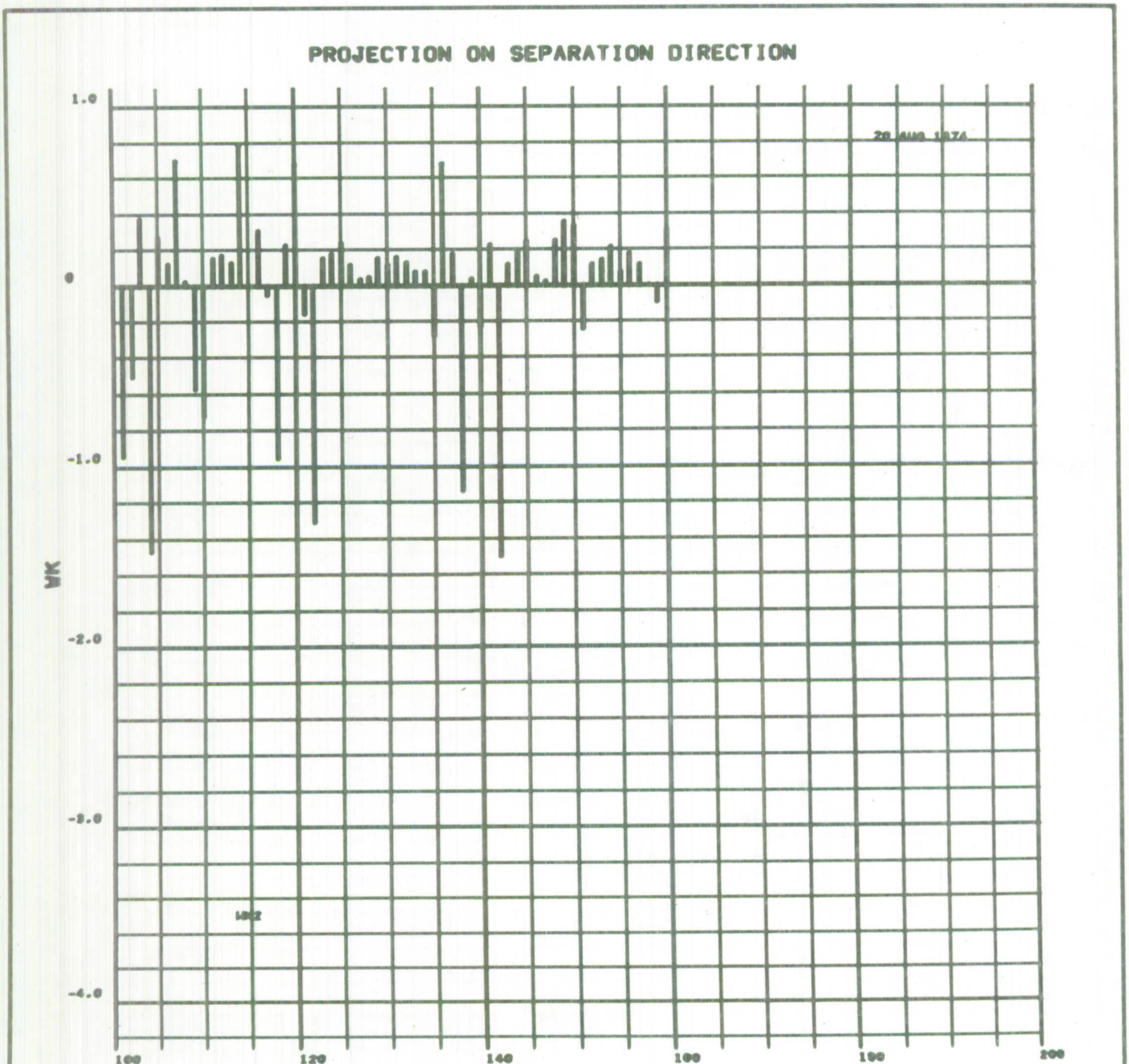


FIGURE 4.4B - BAR CHART SHOWING MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE FISHER CLASSIFICATION DIRECTION USING THE REFERENCE BASE



TARGET 213 ⟶ TARGET 214 ⟶

FIGURE 4.4C - BAR CHART SHOWING MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE FISHER CLASSIFICATION DIRECTION USING THE REFERENCE BASE - TARGET 101

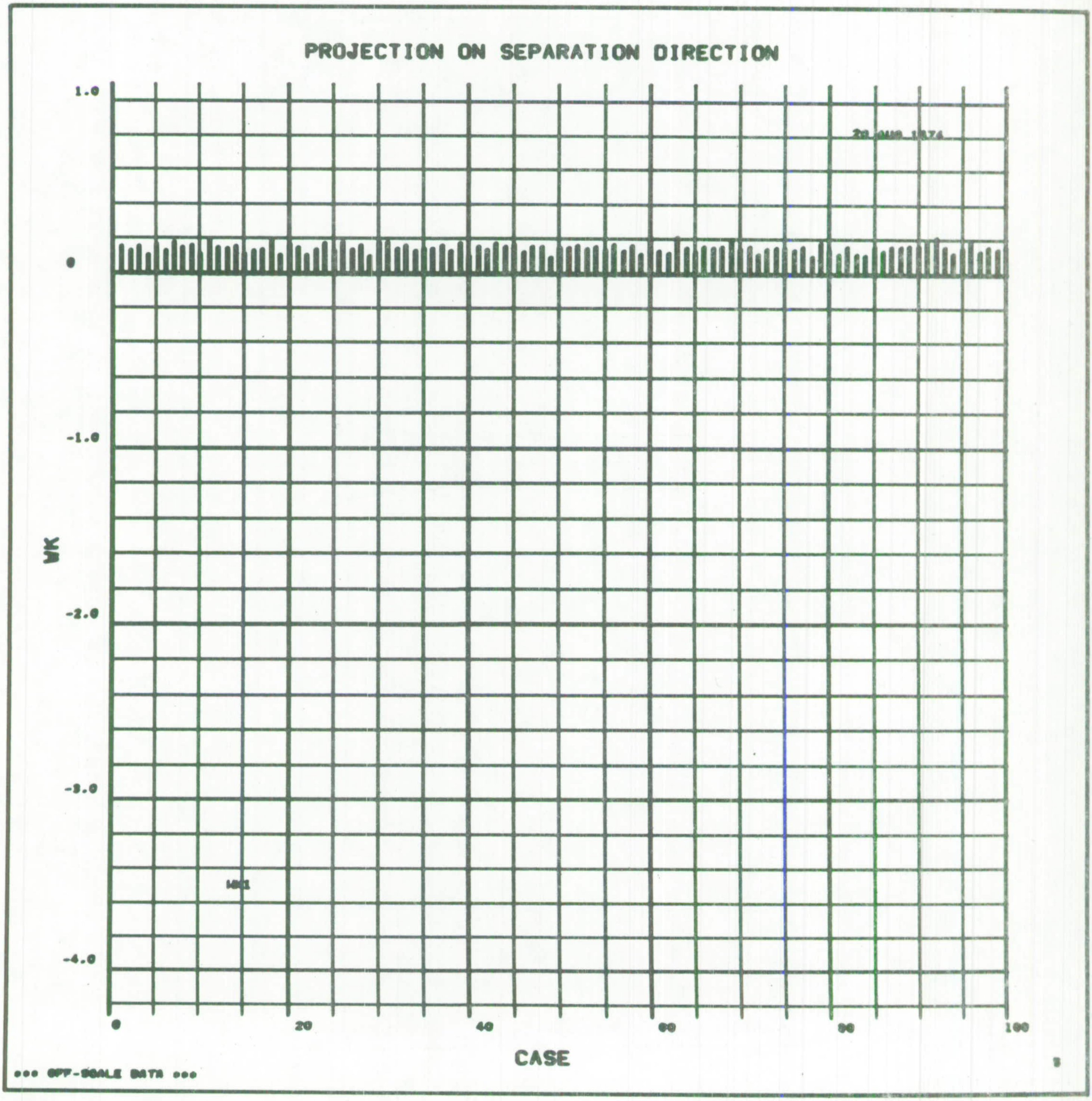
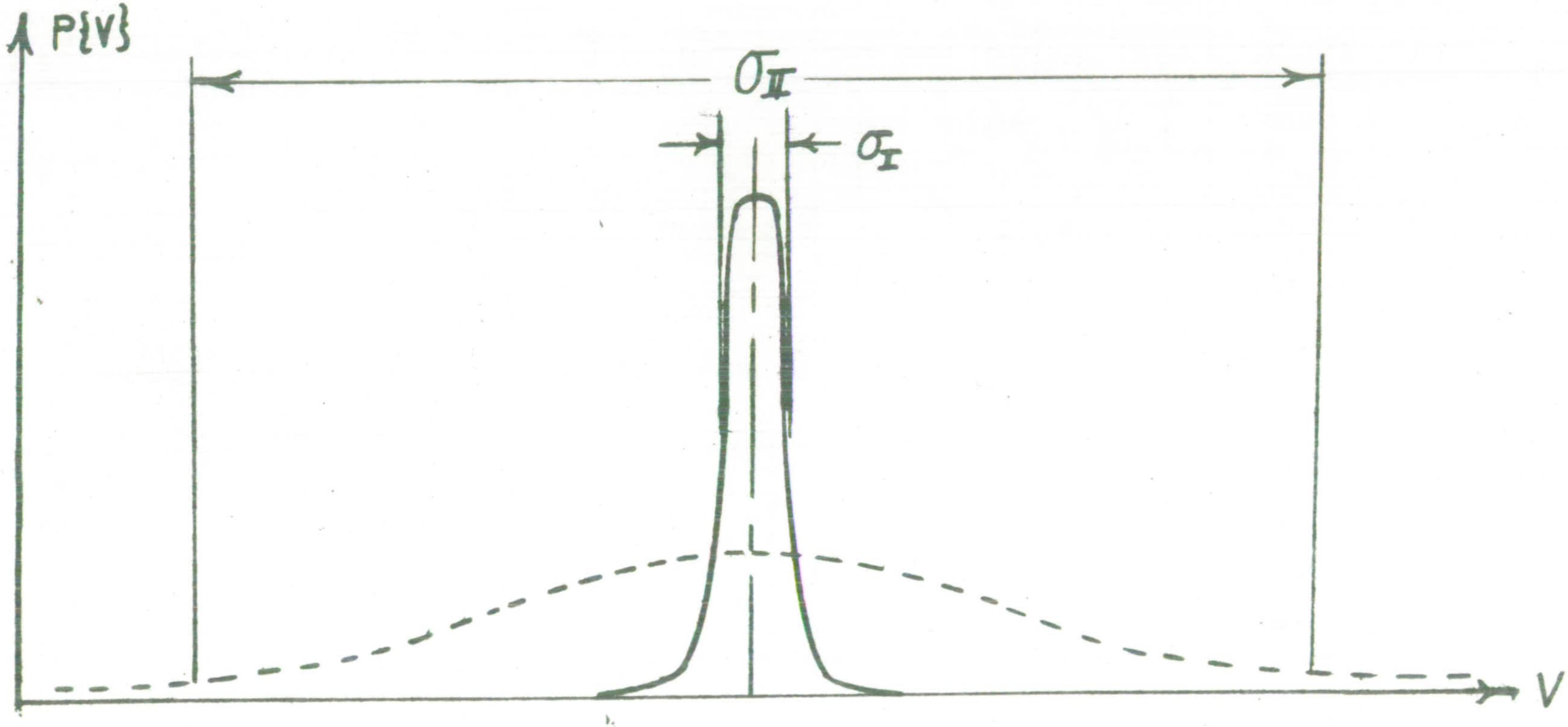


FIGURE 4.5 - LINEAR CLASSIFICATION MECHANISM



the classes. Therefore, several other linear classifiers appear far more satisfactory for this data set.

One linear classifier which is natural for the ADAPT programs is the projection on that direction which contains the greatest amount of variation. By the definition of optimum used in the ADAPT programs, this is just the first optimal coordinate or the projection of the data shown in Figures 4.1 and 4.2 onto the abscissa of these figures. Examination of Figure 4.2 shows that this classification scheme will be quite successful on this learning data since if one used as thresholds for the abscissa of Figure 4.2 the values of -0.32 and -0.2 calling all cases lying between these values members of the 101 Class and all cases lying outside of these values members of the 21X class, only two out of the 260 learning cases would be missed which is considerably better performance than was illustrated by Figures 4.4 even with the double threshold.

It is interesting to note that even the higher order ADAPT optimal coordinates yield good classification results. If one projects on the ordinate of Figure 4.2 and uses thresholds of zero and -0.2 only four of 260 cases are missed and even using the tenth and eleventh optimal directions presented in Figure 4.3 one only misses eight and nine cases out of the 260 learning cases, respectively. This latter phenomenon suggests that one should investigate other techniques than just the projection on the direction of maximum variation since even directions having less variation tend to provide reasonably good separations. One classification scheme that should be better than the Fisher for problems where classes differ primarily in the variation is the classification scheme which we shall call the minimum variation ratio classifier. This classification scheme is a linear classification scheme which selects the direction onto which to project the space by requiring the minimization of the ratio of the variance of one class to the other class.

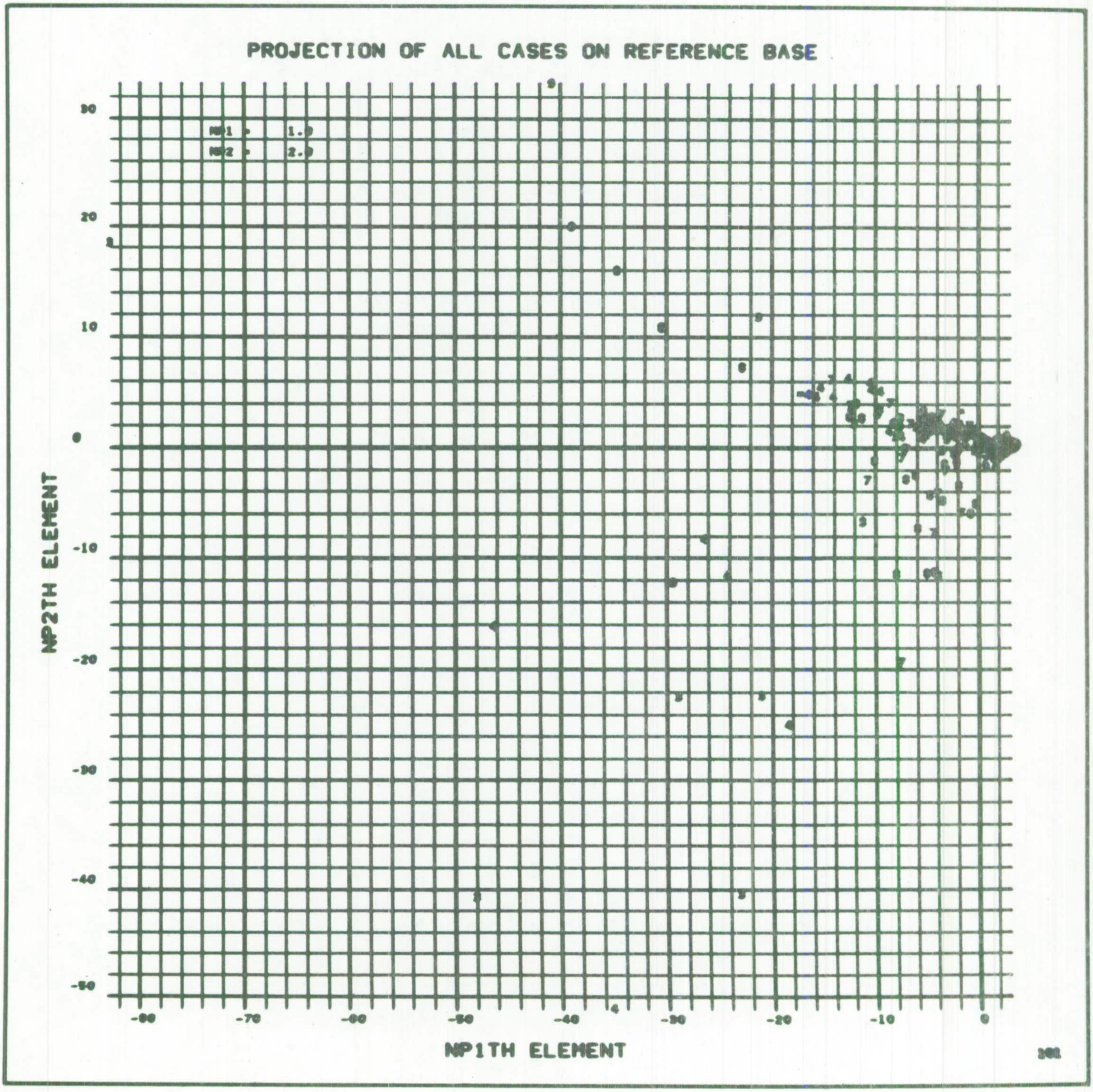
A third solution to the problem of finding an approach to selecting the direction on which to project the space is to non-linearily pre-process the data in such a way that the difference in the variation will be transformed into a difference in the mean values. The simplest such non-linear transformation is simply to square the difference between the location of each point in the space and the centroid of the

class exhibiting the smallest variation. When this is done, the members of the class having the smallest variation will lie near the origin and the members of all the other classes will lie at greater distances from the origin. Thus, the means of the classes can be separated. This is illustrated by Figure 4.6 which is the scatter plot for the squared base. As discussed in Section 3.0, the fact that the centroid of the 101 class is close to zero made it possible to achieve this result by simply squaring the variables. Since the ADAPT representation reduces the noise squaring would probably be most effective if carried out after transformation to the ADAPT coordinate system. However, because of the format of the programs available to perform the analysis, in the present study the squaring was performed on the data prior to the transformation to the ADAPT optimal space. Thus, the third approach to the classification which will be examined in this study is the application of the linear classifier after pre-processing the variables by squaring them.

The character of the data illustrated in the preceding discussion suggests that if the data is ergodic and a high variation case falls within the region defined by the low variation class, this will only occur over a short period of time. If this conjecture is true, then the performance of all of these linear classifiers would be improved by sequential sampling of targets which are not classified as a 21X. If the target is a 21X and time between the samples is large compared to the correlation time of signature, the probability is high that the second data sample will reveal the target to be a 21X rather than a 101 type target. Thus, for any given detection probability, the leakage rate would be improved.

4.2 Performance Comparison of Candidate Linear Classifiers Using Different ADAPT Base

FIGURE 4.6 - SCATTER PLOT OF COEFFICIENTS OF FIRST VERSUS SECOND TERMS IN OPTIMAL REPRESENTATION SHOWING THE LOCATION OF ALL 600 LEARNING AND TEST CASES ON THE SQUARE BASE



Bar charts such as Figures 4.4 are not convenient for comparing a large number of different algorithms each involving a large number of individual cases. Thus, the performance of the various linear classifiers will be presented in two alternate ways. The first is to summarize the characteristics of the detection statistics which would be shown on the bar chart by presenting the: 1) mean value, 2) standard deviation, 3) the max value, 4) the min value and 5) the associated case number for the max and min value of the detection statistics for each class. These tables are only useful for identifying those classifiers which have identical or nearly identical performance. When this occurs, the detection statistic will have the same mean value and standard deviation as well as the same case number and value for the max and min cases.

The second method used here to compare the performance of algorithms is to compare the Receiver Operating Curve, (R.O.C.) or classification performance trade-off curves for the algorithms. These curves consist of the plots of detection probability versus false alarm rate.

Two sets of data were provided for this study. The first consisted of 600 cases divided into 200 members of the 101 Class and 100 members each from Classes 211 through 214. The second set consisted of 1800 cases, 1,000 of which belong to the 101 Class and 200 belong to each of Classes 211 through 214. The smaller of these two sets was supplied initially and was divided into a training and test set of data. The training set consisted of the first hundred of the 101 cases and the first forty of each of the 211 through 214 Classes. The remaining cases were used as a test set. All of the algorithms used in this study were derived using this training set of data. Computationally, the simplest performance evaluation would be to use the performance on the training set. However, this has the disadvantage of the general distrust of the use of learning data to verify the performance of empirical algorithms derived from it and this approach was not considered. Three sets of data were considered for evaluating the performance of the algorithms derived. The first set considered consisted of the 360 case test data set. The second consisted of the entire 600 case set of training plus test data. The first of these suffers from the small number of cases available to develop the classification performance unless one has apriori knowledge of the distribution function. The second set of data

also suffers from the distrust of the use of learning data to verify the performance of empirical algorithms derived from it and also has a relatively small number of samples.

The best set of data for evaluating performance is the 1800 independent test case set. This set of data was not used in the learning and was derived in an identical manner at a different time. It's only disadvantage is that because of the large quantity of cases more computation is required for evaluation especially for non-linear classifiers.

To provide an understanding of the relative merits of each of these data sets for evaluating the performance of the classifiers, the minimum variation ratio algorithm was derived using all three data sets. Since this data was used to construct four different algorithms; namely, the separation of each of Classes 211 through 214 from Class 101, 600, 340 and 1800 case test sets provided 300, 160 and 1200 cases for evaluating each of the algorithms. These evaluations were carried out both under assumption that the detection statistic was Gaussian and under no assumption for the distribution of the detection statistic. When one assumes that the distribution of the detection statistic is Gaussian one may compute the ROC curve from the mean and standard deviation of each of the classes. This curve may be computed for all values desired for the false alarm rate and detection probability. However, extreme care must be used in interpreting these results to assure that one knows the detection statistic is Gaussian even in the wings of the distribution function.

If one does not assume a distribution function for the detection statistic, the ROC curve may still be computed "experimentally" from the values of the detection statistic. This is accomplished by varying the threshold and counting the number of 101 targets detected and the number of 21X targets which have leaked through. By dividing these numbers by the number of cases available in each set one obtains an "experimental" estimate of both the detection probability and the false alarm rate for that particular threshold. The region of the ROC curve which can be computed in this way is limited by the number of test cases. For the 1200 test cases where 1,000 cases were available for the 101 target and 200 for the 21X target one could obtain values for the false alarm rate ranging from .005 to 1 and for the detection probability ranging from 0.001 to 0.999. The values obtained near the outer limits of this region will have significant

uncertainty associated with them. Since this is the only region for which the experimental classification trade-off curve can be obtained, the Gaussian and experimental curves were compared over this region. These comparisons are presented in Figures 4.7 through 4.10 for the minimum variation ratio classifier using the reference base. The solid line on these figures represent the evaluation using the assumption of a Gaussian detection statistic and the 1800 case test data set. The dash line is constructed using the assumption of the Gaussian detection statistic and the 340 case test data set. The dotted line is constructed assuming a Gaussian detection statistic for the 600 cases made up of both the learning and 340 case test data set. The open symbols represent the experimental ROC curve points derived using the 340 case test set. Those open symbols with the flag represent the experimental ROC curve derived using the learning and test data. The closed symbols represent the experimental ROC curve derived using the 1800 case test data set.

Examination of Figures 4.7 through 4.10 shows that for this base and this classification scheme only the experimental ROC curves should be used to obtain estimates of the performance. The ROC curves based on the Gaussian assumptions and the independent test cases consistently overestimate the true performance by approximately the same amount. The 340 independent test sample also underestimates the actual performance even when the experimental data is used. This suggests that there are two components to the error between the Gaussian estimate of 340 test case performance and the experimental performance from the 1800 test cases. One component is due to the non-Gaussian nature of the detection statistic and the other component is due to an actual difference which must exist between these 340 test cases and the 1800 independent test cases. It is recommended that further analysis consisting of developing relative importance vectors between these two data sets be performed to attempt to provide an understanding of whether a real difference does exist and if it does to determine its characteristics.

FIGURE 4.7 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF THE 211 TARGETS ON THE MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE REFERENCE BASE ILLUSTRATING THE RELATIVE EFFECTIVENESS OF SEVERAL EVALUATION SCHEMES

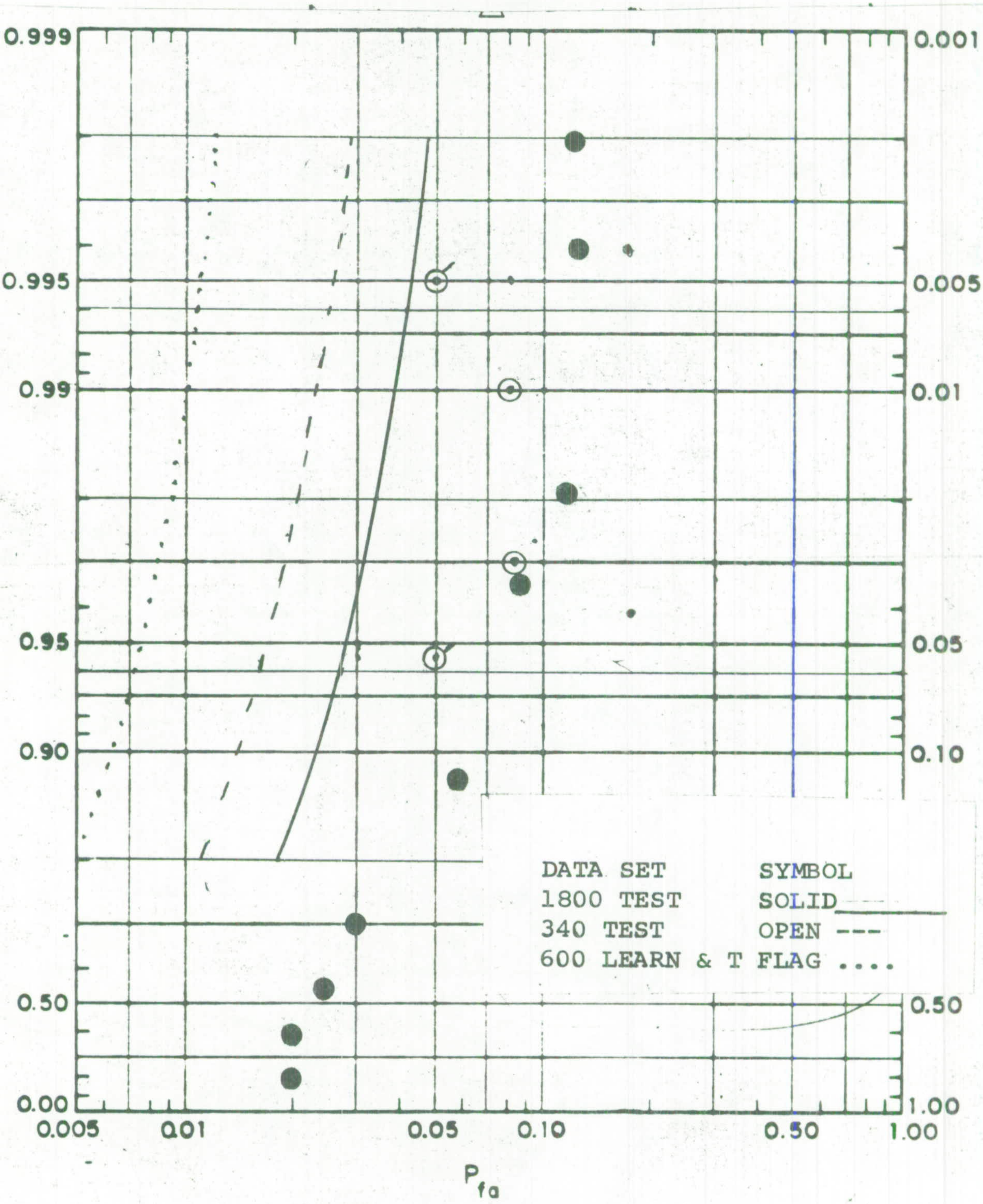


FIGURE 4.8 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF 212 TARGETS ON THE MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE REFERENCE BASE ILLUSTRATING THE RELATIVE EFFECTIVENESS OF SEVERAL EVALUATION SCHEMES

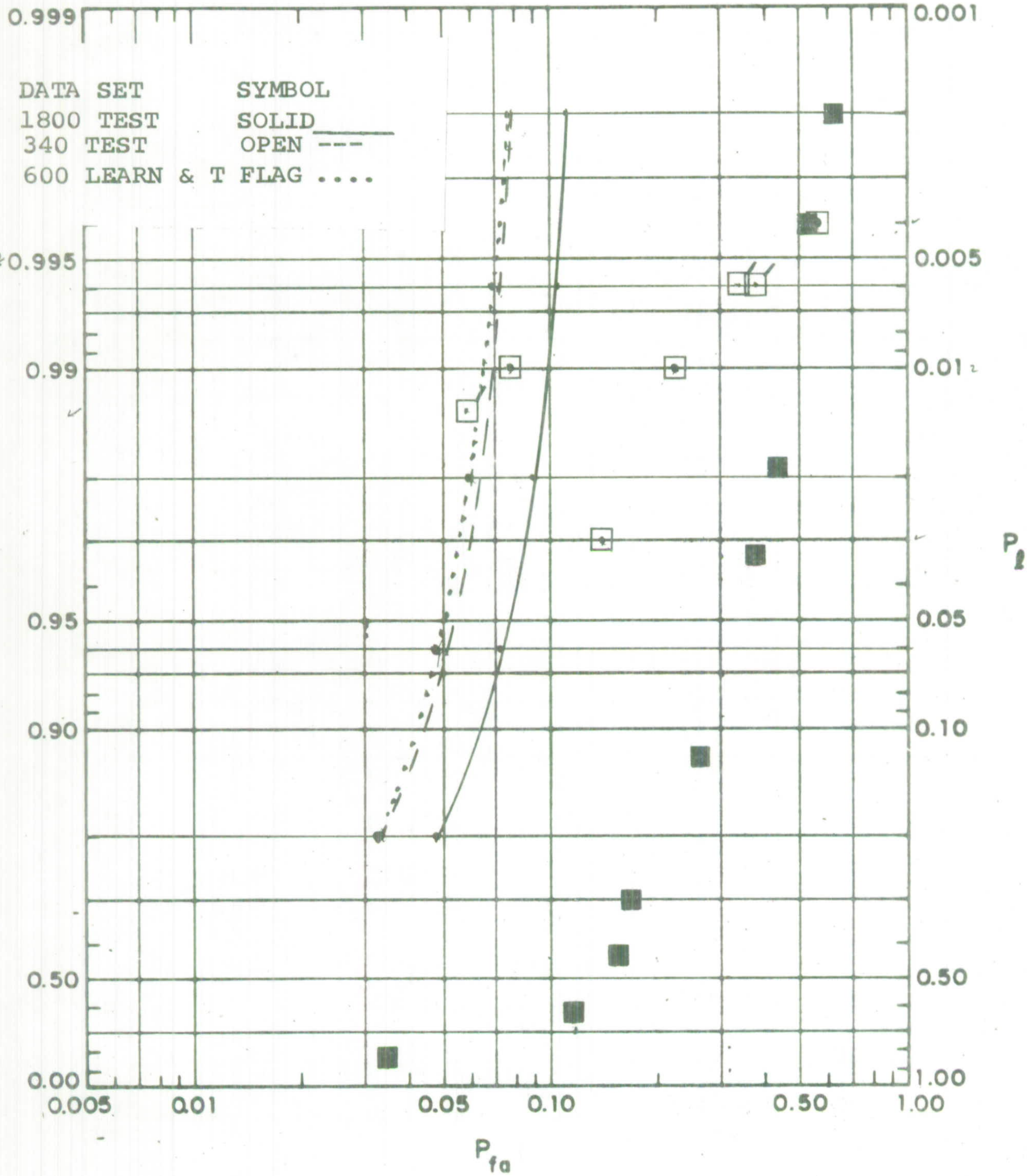


FIGURE 4.9 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF THE 213 TARGETS ON THE MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE REFERENCE BASE ILLUSTRATING THE RELATIVE EFFECTIVENESS OF SEVERAL EVALUATION SCHEMES

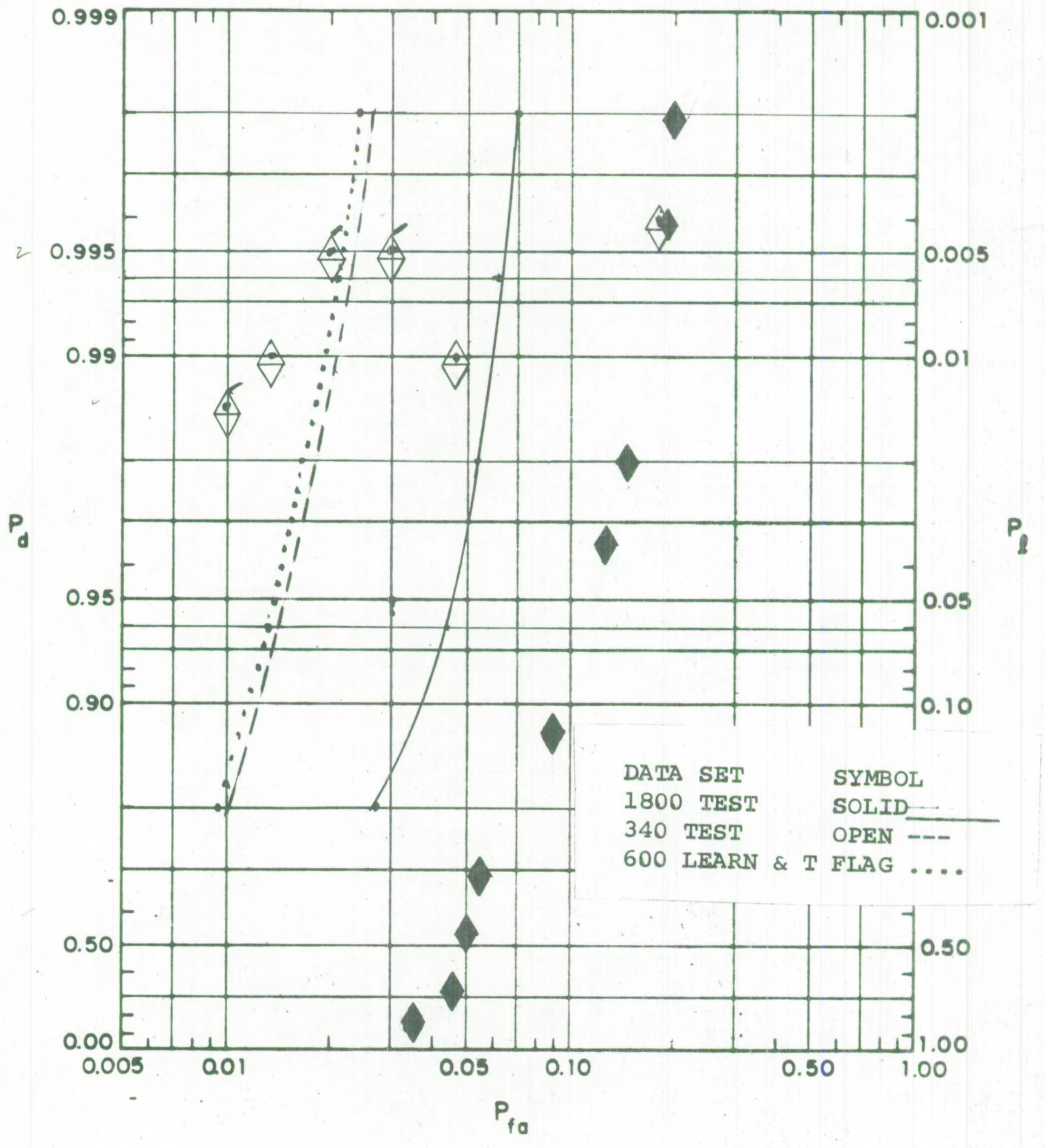
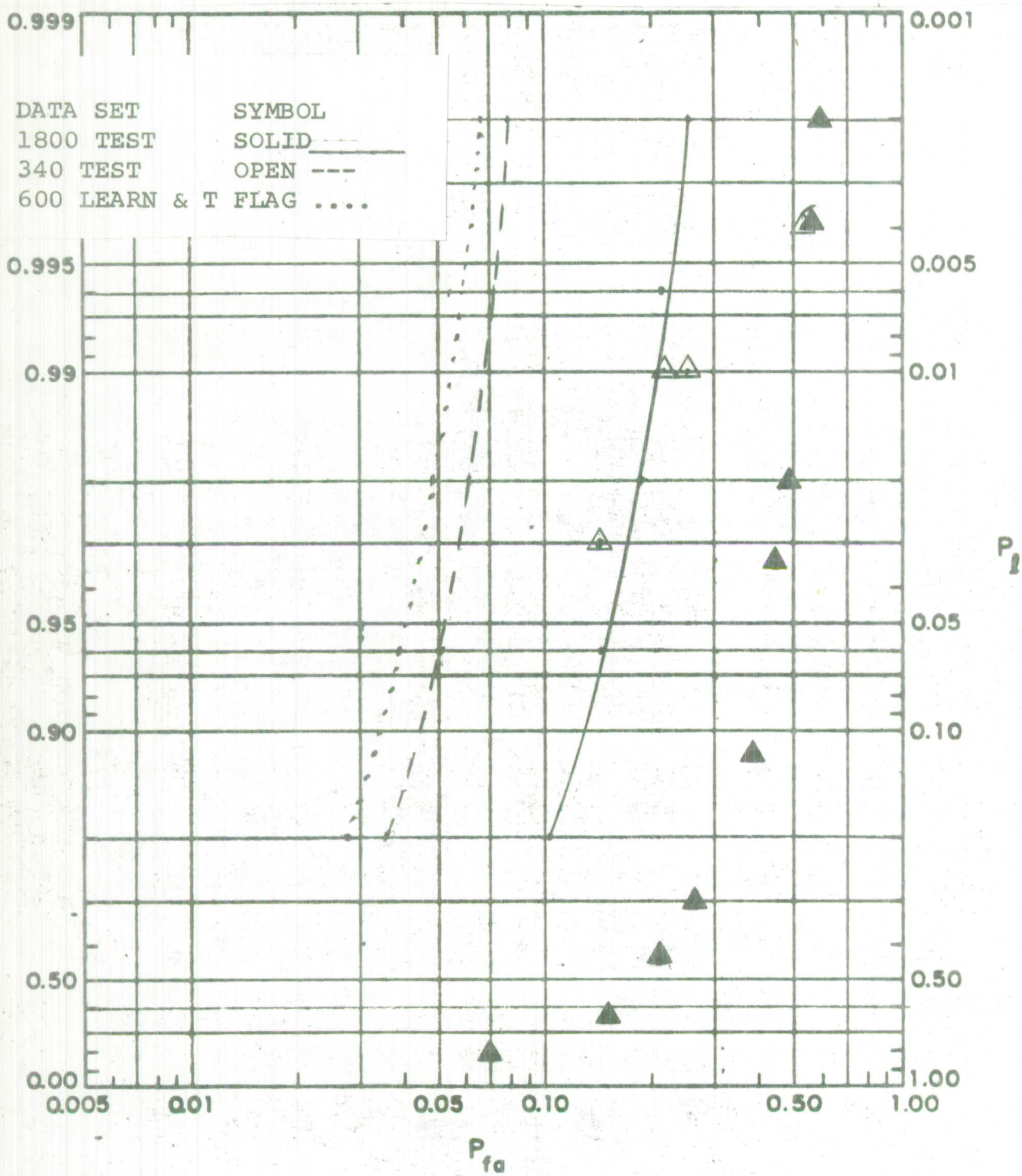


FIGURE 4.10 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF THE 214 TARGETS ON THE MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE REFERENCE BASE ILLUSTRATING THE RELATIVE EFFECTIVENESS OF SEVERAL EVALUATION SCHEMES



4.3 Discrimination Using Projection on First ADAPT Optimal Direction as the Detection Statistic

The projection of the 260 learning cases as well as the 340 independent test cases on the first optimum direction of the reference base is summarized in Table 4.1. This table is typical of the tables used to summarize the projections of this original data set on each of the linear classifiers evaluated in this study. It presents the class identification, the number of members in each class, the mean value, the standard deviation, max and min values of the projection and identifies the case or count number within each class at which the max and min values occur. The classes are identified in Column 1 of Tables 4.1 through 4.10 by the Numerals 1-10 where the Numeral "1" identifies the 100 independent test cases belonging to Class 101, even Numerals "2" through "8" identify the learning cases for Classes 211, 212, 213 and 214, respectively. Numeral "10" identifies the learning cases for the 101 Class and the odd Numerals 3 through 9 identify the 60 independent test cases belonging to Classes 211, 212, 213 and 214, respectively. Thus, Table 4.1 shows that the 40 learning cases used for the 212 Class had a mean value for the projection on the first ADAPT optimal direction of -0.261 with a standard deviation of 0.418. The max and min values were 1.032 and -0.725, respectively, and occurred at the 27th and 37th case, respectively. Similarly, the mean value of the projection of the corresponding 60 independent test cases on the first ADAPT optimal coordinate is -0.313 with a standard deviation of 0.561. Max and min values for the independent test data occurred at the 2nd and 9th case, respectively, and had values of 1.67 and -2.4, respectively.

TABLE 4.1 - DETECTION STATISTIC SUMMARY FOR FIRST ADAPT OPTIMAL DIRECTION CLASSIFIER USING REFERENCE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	-0.2570E 00	0.2450E-01	-0.2127E 00	80	-0.3455E 00	62
2	40	0.8433E 00	0.4430E 01	0.1621E 02	2	-0.1417E 02	25
3	60	-0.2502E 00	0.1759E 01	0.7421E 01	9	-0.3474E 01	41
4	40	-0.2610E 00	0.4176E 00	0.1032E 01	27	-0.7252E 00	37
5	60	-0.3126E 00	0.5607E 00	0.1674E 01	2	-0.2403E 01	9
6	40	0.1740E 00	0.1801E 01	0.4554E 01	6	-0.4474E 01	15
7	60	0.2780E 00	0.2019E 01	0.7612E 01	40	-0.3708E 01	58
8	40	-0.1025E 00	0.1031E 01	0.5582E 01	22	-0.1539E 01	1
9	60	-0.2515E 00	0.7460E 00	0.4045E 01	13	-0.1452E 01	49
10	100	-0.2615E 00	0.2246E-01	-0.2118E 00	44	-0.3095E 00	52

TABLE 4.2 - DETECTION STATISTIC SUMMARY FOR MINIMUM VARIATION RATIO CLASSIFIER USING THE REFERENCE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	-0.1381E-01	0.9022E-02	0.1428E-01	62	-0.3400E-01	64
2	40	-0.2148E 00	0.1872E 01	0.8294E 01	25	-0.5470E 01	10
3	60	-0.2495E-01	0.6314E 00	0.1353E 01	40	-0.3174E 01	9
4	40	0.3489E-02	0.1525E 00	0.4926E 00	16	-0.3047E 00	40
5	60	-0.2538E-02	C.2119E 00	0.5284E 00	40	-0.5838E 00	29
6	40	-0.1162E 00	0.6125E 00	0.1537E 01	12	-0.2218E 01	30
7	60	-0.1368E 00	0.6701E 00	0.1147E 01	2	-0.2751E 01	40
8	40	-0.4199E-01	0.2613E 00	0.3549E 00	1	-0.1372E 01	22
9	60	-0.1444E-01	0.2112E 00	0.4748E 00	12	-0.9861E 00	13
10	100	-0.1182E-01	0.6146E-02	0.5088E-02	12	-0.2625E-01	51

TABLE 4.3 - DETECTION STATISTIC SUMMARY FOR MINIMUM VARIATION RATIO CLASSIFIER USING THE TIME BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	-0.1376E-01	0.9020E-02	0.1429E-01	62	-0.3392E-01	64
2	40	-0.2171E 00	0.1870E 01	0.8272E 01	25	-0.5461E 01	10
3	60	-0.2557E-01	0.6313E 00	0.1347E 01	40	-0.3174E C1	9
4	40	0.3360E-02	0.1531E 00	0.4937E 00	16	-0.3077E C0	40
5	60	-0.2753E-02	0.2123E 00	0.5267E 00	40	-0.5851E 00	29
6	40	-0.1177E 00	0.6140E 00	0.1526E 01	12	-0.2241E 01	30
7	60	-0.1356E 00	0.6691E 00	0.1137E 01	2	-0.2745E 01	40
8	40	-0.4184E-01	0.2617E 00	0.3612E 00	1	-0.1374E 01	22
9	60	-0.1485E-01	0.2120E 00	0.4730E 00	12	-0.1004E 01	13
10	100	-0.1178E-01	0.6143E-02	0.5108E-02	12	-0.2622E-01	51

TABLE 4.4 - DETECTION STATISTIC SUMMARY FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE PRINCIPAL POLARIZATION TIME BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.1691E-01	0.1173E-01	0.3692E-01	83	-0.4546E-01	62
2	40	0.5436E 00	0.1703E 01	0.6480E 01	2	-0.3980E 01	25
3	60	-0.1105E-01	0.9152E 00	0.3085E 01	16	-0.2176E C1	40
4	40	0.6631E-02	0.2103E 00	0.5963E 00	18	-0.3197E 00	16
5	60	-0.1644E-01	0.2947E 00	0.6558E 00	29	-0.1553E 01	9
6	40	0.1801E 00	0.8261E 00	0.2703E 01	24	-0.1391E 01	15
7	60	0.2578E 00	0.9965E 0C	0.4510E 01	40	-0.1286E 01	31
8	40	0.1726E 00	0.4674E 00	0.2614E 01	22	-0.1583E 00	29
9	60	0.2055E-01	0.3105E 0C	0.1658E 01	41	-0.4073E 00	49
10	100	0.1736E-01	0.7355E-02	0.4191E-01	93	-0.1191E-01	37

TABLE 4.5 - DETECTION STATISTIC SUMMARY FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE PRINCIPAL POLARIZATION BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.1691E-01	0.1173E-01	0.3692E-01	83	-0.4546E-01	62
2	40	0.5436E 00	0.1703E 01	0.6480E 01	2	-0.3980E 01	25
3	60	-0.1105E-01	0.9152E 00	0.3085E 01	16	-0.2176E 01	40
4	40	0.6631E-02	0.2103E 00	0.5963E 00	18	-0.3197E 00	16
5	60	-0.1644E-01	0.2947E 00	0.6558E 00	29	-0.1553E 01	9
6	40	0.1801E 00	0.8261E 00	0.2703E 01	24	-0.1391E 01	15
7	60	0.2578E 00	0.9964E 00	0.4510E 01	40	-0.1286E 01	31
8	40	0.1726E 00	0.4674E 00	0.2614E 01	22	-0.1583E 00	29
9	60	0.2055E-01	0.3105E 00	0.1658E 01	41	-0.4073E 00	49
10	100	0.1736E-01	0.7355E-02	0.4191E-01	93	-0.1191E-01	37

TABLE 4.6 - DETECTION STATISTIC SUMMARY FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE CLASS II ONLY BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.1380E-01	0.9022E-02	0.3408E-01	64	-0.1422E-01	62
2	40	0.2149E 00	0.1872E 01	0.5466E 01	10	-0.8295E 01	25
3	60	0.2520E-01	0.6316E 00	0.3176E 01	9	-0.1354E 01	40
4	40	-0.3361E-02	0.1527E 00	0.3057E 00	40	-0.4931E 00	16
5	60	0.2545E-02	0.2120E 00	0.5843E 00	29	-0.5279E 00	40
6	40	0.1168E 00	0.6131E 00	0.2224E 01	30	-0.1535E 01	12
7	60	0.1370E 00	0.6699E 00	0.2751E 01	40	-0.1148E 01	2
8	40	0.4203E-01	0.2612E 00	0.1372E 01	22	-0.3531E 00	1
9	60	0.1461E-01	0.2116E 00	0.9870E 00	13	-0.4747E 00	12
10	100	0.1182E-01	0.6147E-02	0.2625E-01	51	-0.5097E-02	12

TABLE 4.7 - DETECTION STATISTIC SUMMARY FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE 3-CLASS II TARGET BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	-0.1417E-01	0.8010E-02	0.8473E-02	40	-0.3223E-01	64
2	40	-0.2165E 00	0.1628E 01	0.7203E 01	25	-0.4808E 01	10
3	60	-0.2750E-01	0.5680E 00	0.1034E 01	40	-0.2587E 01	9
4	40	-0.6976E-03	0.1441E 00	0.4351E 00	16	-0.2841E 00	27
5	60	-0.2030E-02	0.1966E 00	0.4736E 00	6 40	-0.5513E 00	14
6	40	-0.1099E 00	0.5522E 00	0.1288E 01	12	-0.1985E 01	30
7	60	-0.1406E 00	0.6026E 00	0.1118E 01	2	-0.2510E 01	40
8	40	-0.1728E-01	0.1921E 00	0.3869E 00	1	-0.8875E 00	22
9	60	-0.1934E-01	0.1984E 00	0.4339E 00	12	-0.8594E 00	13
10	100	-0.1274E-01	0.5507E-02	0.1850E-02	12	-0.2413E-01	10

TABLE 4.8 - DETECTION STATISTIC SUMMARY FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE SQUARE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.1215E-03	0.7253E-04	0.4456E-03	62	-0.2073E-04	96
2	40	0.6817E 00	0.1872E 01	0.9068E 01	25	0.8350E-03	1
3	60	0.4008E 00	0.1262E 01	0.7840E 01	9	-0.9306E-01	8
4	40	0.1850E-01	0.2342E-01	0.8632E-01	27	0.1168E-02	29
5	60	0.2475E-01	0.3349E-01	0.1473E 00	30	0.1290E-02	59
6	40	0.2937E 00	0.4518E 00	0.2641E 01	2	-0.5535E-01	20
7	60	0.3839E 00	0.5910E 00	0.3163E 01	51	0.1839E-03	55
8	40	0.3866E-01	0.7254E-01	0.3845E 00	22	0.3069E-03	37
9	60	0.3679E-01	0.7630E-01	0.4355E 00	41	-0.5143E-02	59
10	100	0.1173E-03	0.5305E-04	0.2753E-03	12	0.5405E-05	11

TABLE 4.9 - DETECTION STATISTIC SUMMARY FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE ZERO MEAN SQUARE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.1165E-03	0.7253E-04	0.4417E-03	62	-0.2463E-04	96
2	40	0.6763E 00	0.1867E 01	0.8933E 01	25	-0.1038E-01	1
3	60	0.3965E 00	0.1257E 01	0.7789E 01	9	-0.1004E 00	8
4	40	0.1822E-01	0.2304E-01	0.7993E-01	27	0.1125E-02	29
5	60	0.2419E-01	0.3267E-01	0.1406E 00	14	0.1150E-C2	18
6	40	0.2923E 00	0.4545E 00	0.2666E 01	2	-0.7046E-01	20
7	60	0.3815E 00	0.5944E 00	0.3158E 01	51	-0.3664E-02	55
8	40	0.3880E-01	0.7403E-01	0.3940E 00	22	0.3043E-03	37
9	60	0.3630E-01	0.7554E-01	0.4300E 00	41	-0.5537E-02	59
10	100	0.1123E-03	0.5282E-04	0.2681E-03	12	-0.4023E-05	11

TABLE 4.10 - DETECTION STATISTIC SUMMARY FOR FIRST ADAPT OPTIMAL DIRECTION CLASSIFIER DEVELOPED USING THE SQUARE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.2975E 01	0.1894E-01	0.2979E 01	73	0.2964E 01	62
2	40	-0.9309E 01	0.1733E 02	0.2896E 01	13	-0.8297E 02	2
3	60	-0.5490E 01	0.9255E C1	0.2716E 01	23	-0.4619E 02	26
4	40	0.2558E 01	0.1050E C1	0.2951E 01	24	-0.3762E 01	27
5	60	0.2089E 01	0.2149E 01	0.2952E 01	39	-0.7775E 01	14
6	40	-0.1699E 01	0.4590E 01	C.2722E 01	31	-0.1419E 02	20
7	60	-0.2736E 01	0.6373E 01	0.2844E 01	9	-0.3041E 02	40
8	40	0.1014E 01	0.6932E 01	0.2977E 01	17	-0.4074E 02	22
9	60	0.1975E 01	0.2882E 01	0.2964E 01	9	-0.1713E 02	41
1C	100	0.2975E 01	0.1868E-01	0.2978E 01	67	0.2967E 01	82

The performance of this classifier on the 1800 independent test cases is summarized by the four classification trade-off curves presented in Figure 4.11. This curve shows the detection probability versus false alarm rate for each of the 21X targets when the first ADAPT optimal coefficient is used as the detection statistic for separating the 21X target from the 101 target. These curves show that there is considerable variation in the performance between the various 21X targets. For example, if we desire detection probability for the 101 target of 0.99, the false alarm rate for the 211 target will be between 3 and 4% and the false alarm rate for the 214 target will be almost 50%.

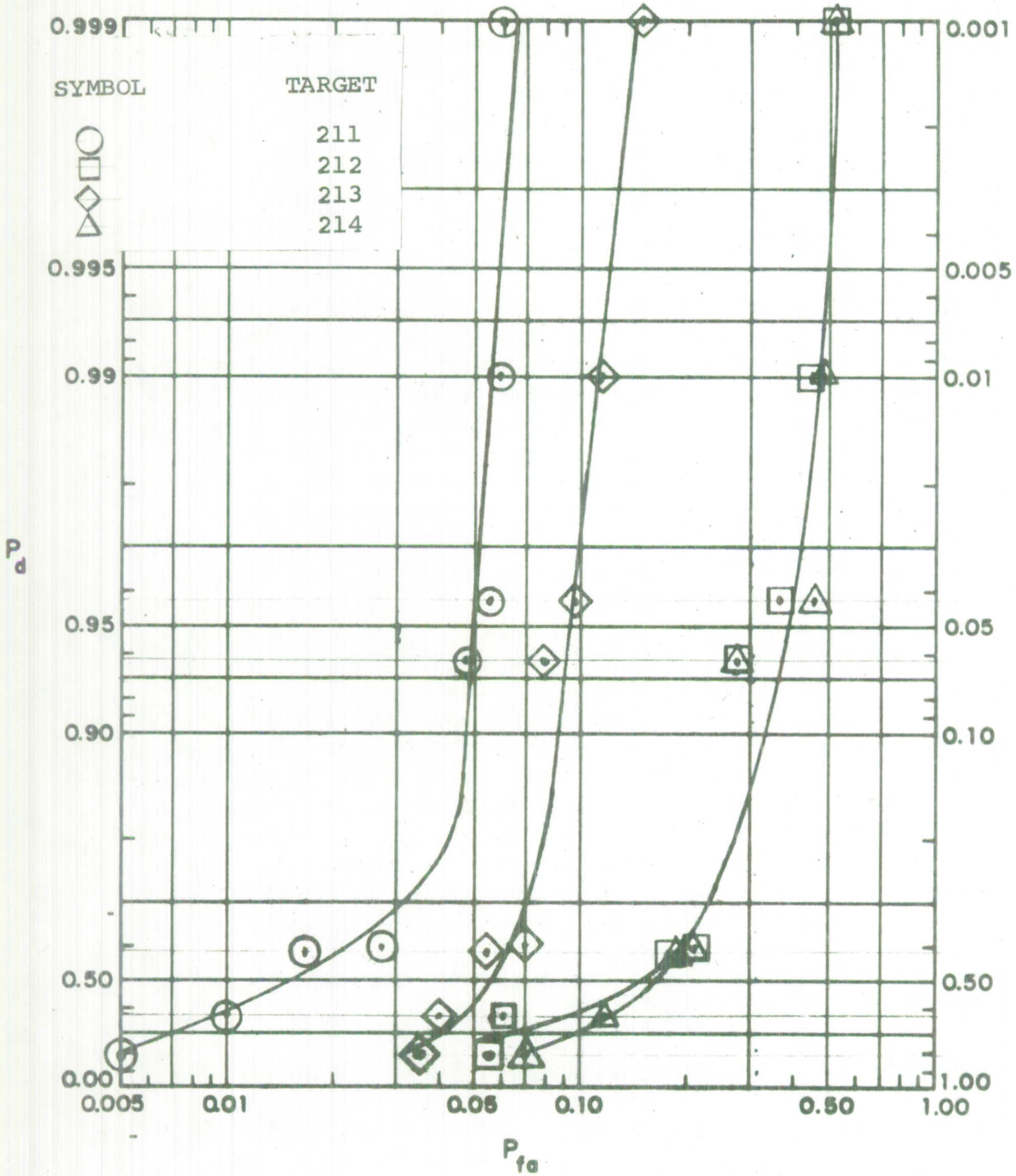
The relative importance vector for this classifier is of course simply the first optimal function of the reference base which was presented in Figure 3.5. The double thresholding(1) of the detection statistic makes the interpretation of the relative importance vector different than the interpretation for a single threshold classifier such as the Fisher discriminant used in Reference 1. The absolute magnitude of the relative importance vector for any given indexing variable still provides an indication of how important that variable is to reaching the decision. However, since both large and small values of the detection statistic now belong to the same class it is considerably more difficult to associate a given feature of the relative importance vector with some characteristics of one of the classes. Examination of Figure 3.5 shows the surprising result that for this classification scheme, the time portion of the data history (indexing variables 1 through 20 and 41 through 60) appeared to be significantly more important than the frequency domain. Since both the linearity of the Fourier transform and the analysis of the reference base as compared to the time only base show these two bases to contain identically the same information, it is difficult to understand why the time domain should be more useful than the frequency domain. This phenomena is also discussed in Section 3.4.

4.4 Discrimination Using the Minimum Variation Ratio Classifier

The minimum variation ratio classifier is a classifier which projects all of the data on that direction which

(1) Double thresholding consists of requiring the detection statistic to lie between two numbers.

FIGURE 4.11 - CLASSIFICATION TRADE-OFF CURVE FOR PROJECTION OF ALL TARGETS ON THE FIRST ADAPT OPTIMAL DIRECTION CLASSIFIER DEVELOPED USING THE REFERENCE BASE



minimizes the ratio of the variation in one class to the variation in the other class. This classifier is also known as Simultaneous Diagonalization. Intuitively, one expects that this should be a very good criteria for data which has the spatial distribution observed in the scatter plots. Since it examines the entire optimal space it may yield a better classification direction then projecting on the first ADAPT optimal coordinate. Since this classifier is more representative of other linear classifiers which might be evaluated using the ADAPT optimal representation, it was selected as the reference classifier for the present study.

The projection of the 260 learning cases on the minimum variation ratio classifier optimal direction is illustrated in Figures 4.12A through 4.12C. Figures 4.12A and B show the projection of the 21X class on this classifier and Figure 4.12C showed a projection of the 101 Class on this classifier. This particular classifier was derived to minimize the ratio of the variation in the 101 Class to the variation of the 21X Class. The information presented on Figure 4.12 may be summarized in a table such as Table 4.2. Table 4.2 summarizes this projection both for the learning data shown in Figure 4.12 and the 340 independent test cases. The class identification and values presented are the same as those presented in Tables 4.1. The mean value for the learning 101 Class is -0.01182. Because of a difference in the origin used in the computer programs which prepared Table 4.2 and Figure 4.12 and additive constant of .0614 must be added to the values given in the table to obtain the corresponding value on Figure 4.12. Thus, this value of -0.0118 corresponds to the value of .0496 which is plotted in Figure 4.12C. Because of the small value of the standard deviation of this class, it is difficult to visually verify that case numbers 12 and 51 on Figure 4.12C are the maximum and minimum values of the detection statistic, respectively. However, examination of the first 40 points on Figure 4.12A allows one to easily verify that the 25th case having a value of approximately 8.8 is the maximum value and that the 10th case having a value approximately -4.86 (i.e. -5.47 plus the additive constant of .614) is the minimum value for the 211 class. The reader may make similar comparisons for the 212, 213 and 214 classes as presented in Figure 4.12A.

The relative importance spectrum for the minimum variation ratio classifier is presented in Figure 4.13. This bar chart shows the relative importance of each of the ADAPT optimal

FIGURE 4.12A - BAR CHART SHOWING THE MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION USING THE REFERENCE BASE

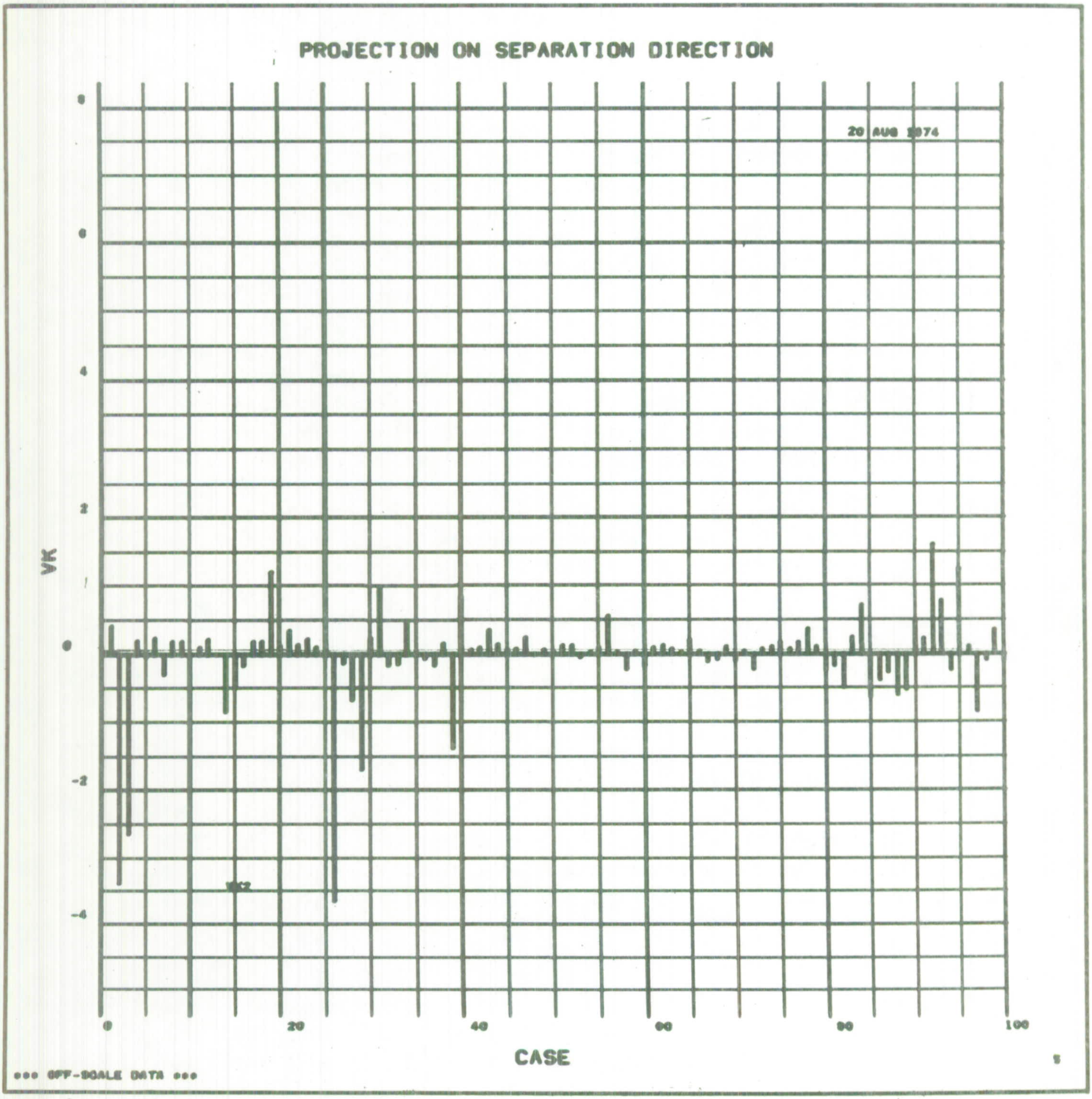


FIGURE 4.12B - BAR CHART SHOWING THE MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION USING THE REFERENCE BASE

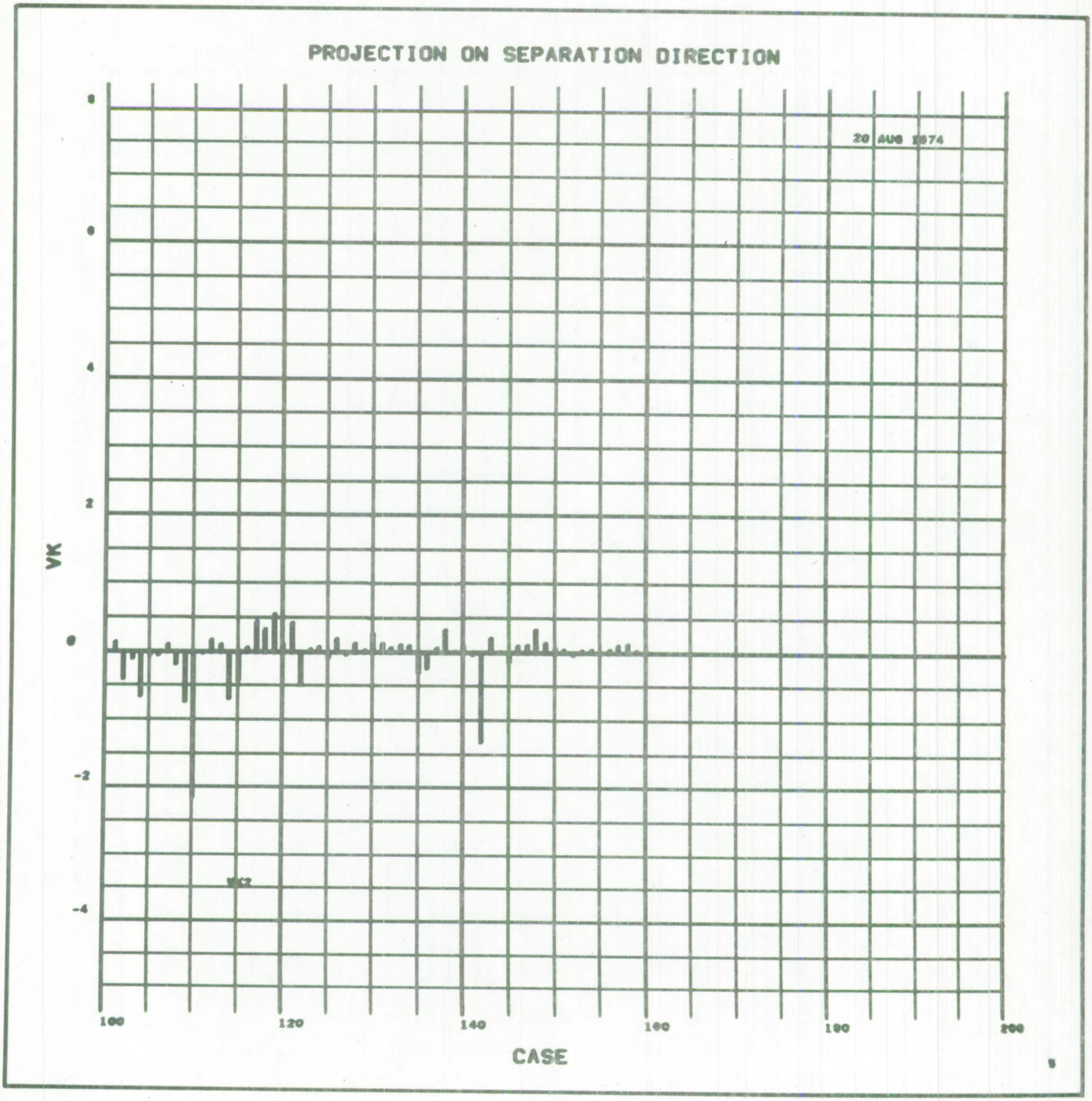


FIGURE 412C - BAR CHART SHOWING THE MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION USING THE REFERENCE BASE FOR TARGET 101

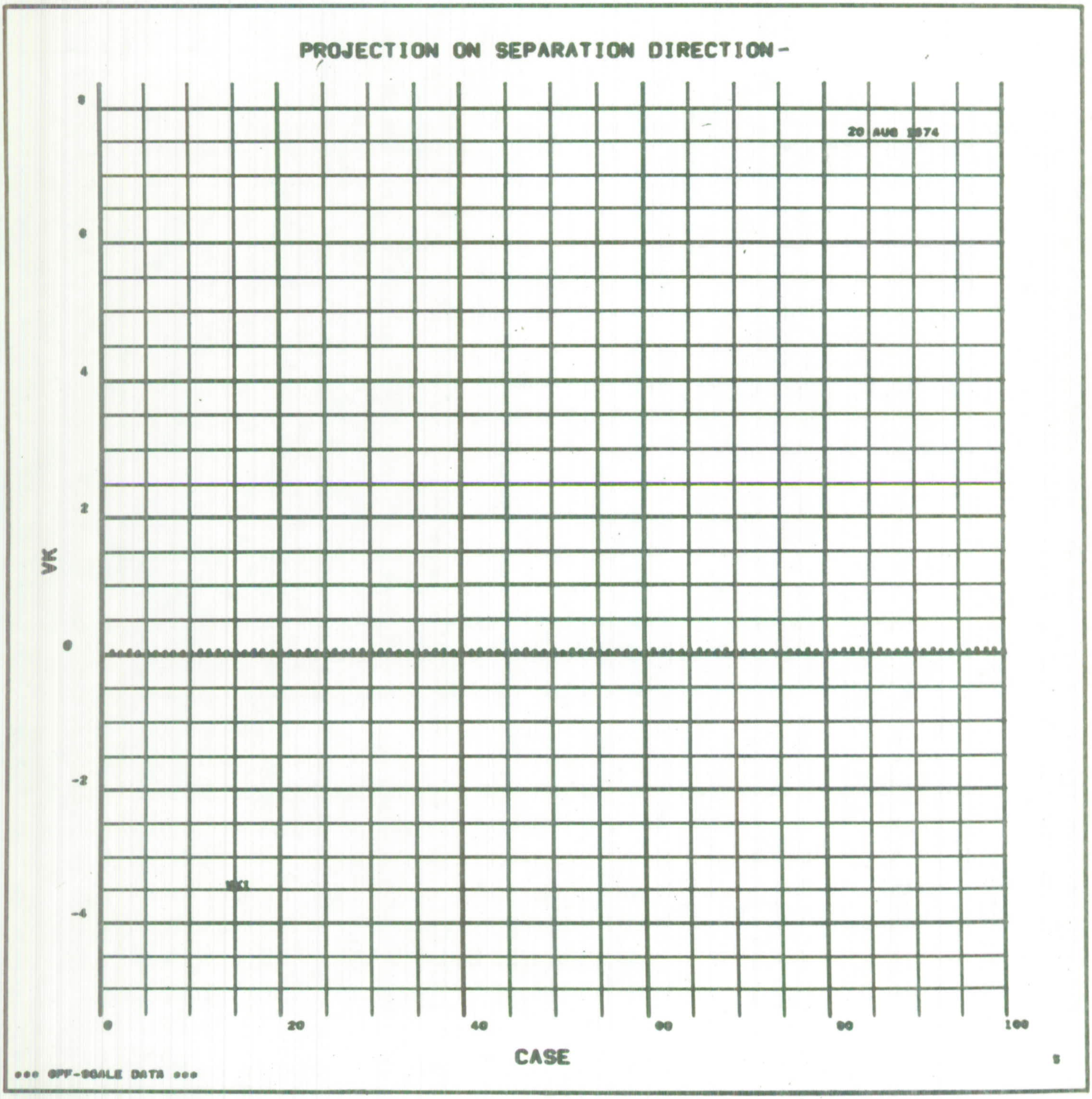
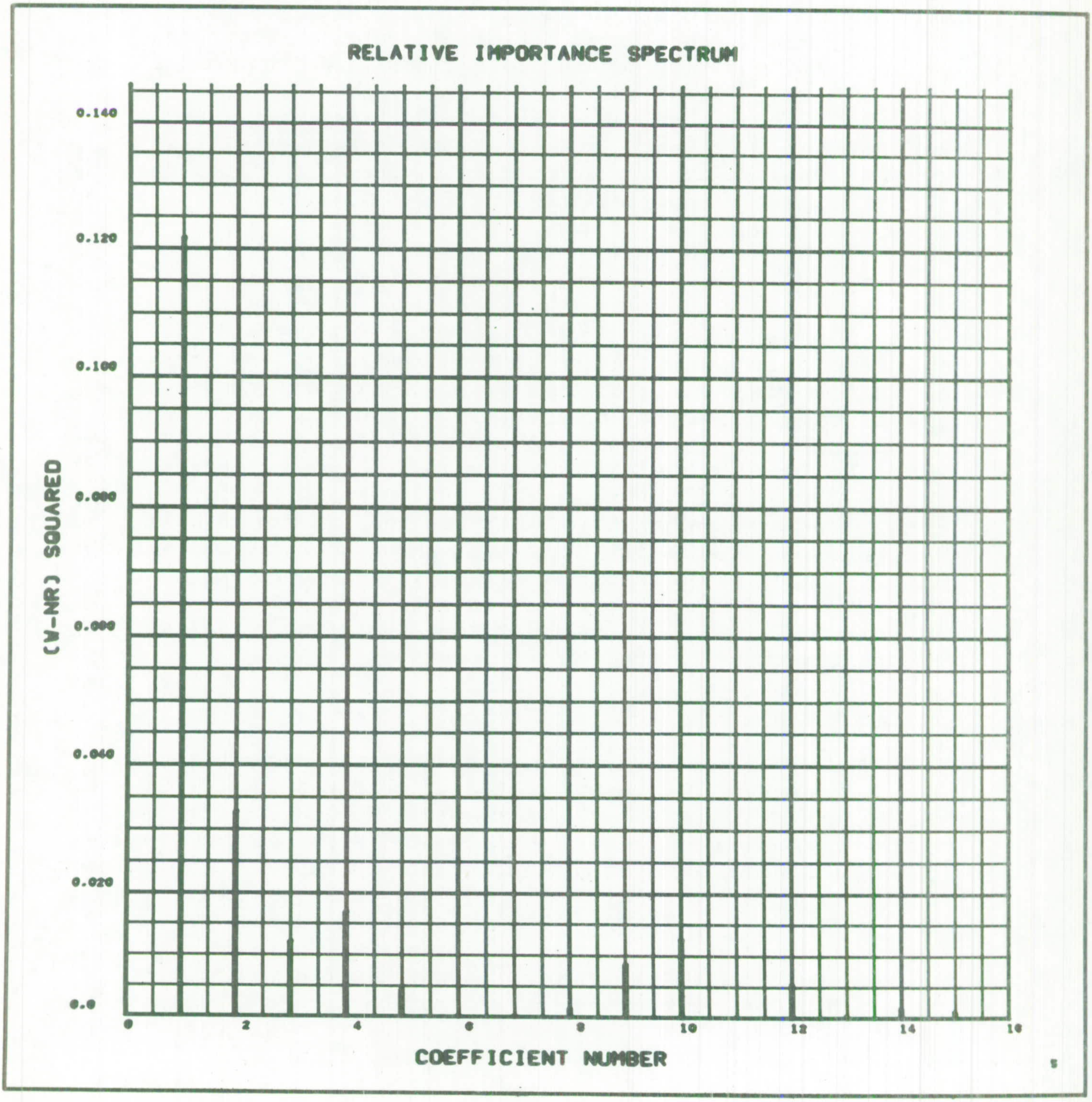


FIGURE 4.13 - RELATIVE IMPORTANCE OF OPTIMAL COORDINATES TO MINIMUM VARIATION CLASSIFIER DEVELOPED USING THE REFERENCE BASE



coordinate directions to the direction on which all of the data is being projected. This figure shows that the minimum variation ratio classifier used here is primarily lined up with the first optimal coordinate but is slightly deflected towards the direction of the second, fourth and tenth optimal coordinates. The projection of this direction onto the original data space gives the relative importance vector for this classifier which is presented in Figure 4.14. Examination of this figure leads to the same major conclusion as examination of the first optimal function in Figure 3.5; namely, that the time domain plays a far more important role in the classification then the frequency domain. The detailed structure of the relative importance vector shows considerable difference from the detail structure of the first optimal function presented in Figure 3.5. Thus, this is a different classifier then the first ADAPT optimal direction classifier.

The performance of this classifier for separating the four 21X targets from the 101 target is summarized in the classification trade-off curve presented in Figure 4.15. Figure 4.15 is a compilation of the experimental ROC curves taken from Figures 4.7 through 4.10. This curve shows that the relative difficulty of identifying the four 21X targets is approximately in the same order as for the first ADAPT optimal direction classifier. However, the total variation between the easiest and most difficult classifiers is considerably less than that which was seen in Figure 4.11. Comparison of Figure 4.15 with 4.11 shows that the classification performance of these two classifiers is quite similar for the more difficult targets, but the first ADAPT optimal direction classifier is superior to the minimum variation ratio classifier for the easier targets.

The Effect of Radar System Complexity

The reference base was used to evaluate the effect of the radar system complexity on the classification performance. The effect of the radar system complexity was evaluated by comparing the performance with similar bases and classification schemes derived utilizing only the time domain portion of the signature and only the principal polarization. This allows the estimate of the degradation in performance that occurs if only the principal polarization is measured and the degradation in performance which would be incurred if one did not perform the transformation to the frequency domain.

FIGURE 4.14 - RELATIVE IMPORTANCE OF SIGNAL ELEMENT CORRESPONDING TO INDEXING VARIABLE FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE REFERENCE BASE

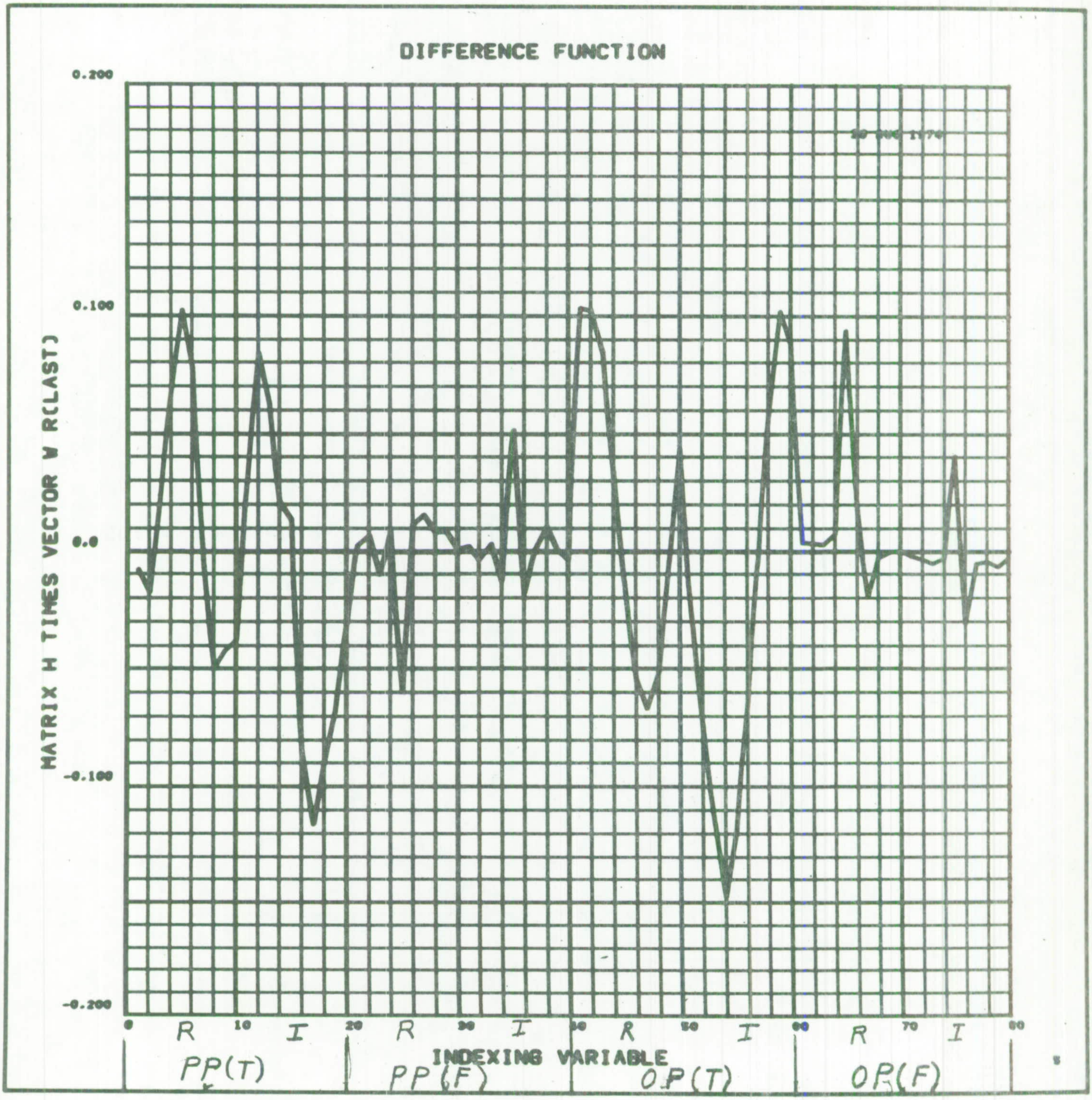


FIGURE 4.15 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF ALL THE TARGETS ON THE MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE REFERENCE BASE

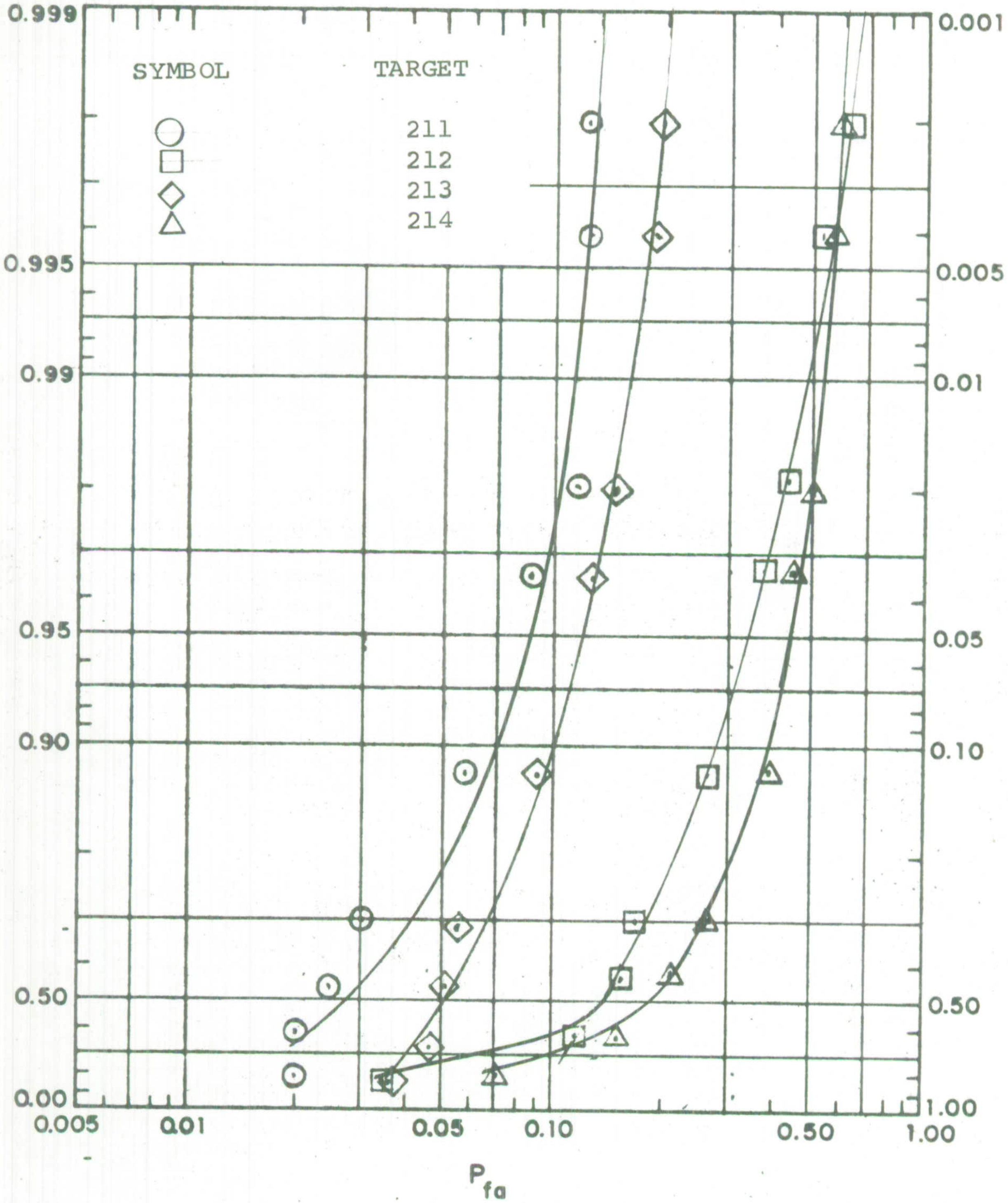


Table 4.3 summarizes the projection of the 260 learning cases and 340 independent test cases on the minimum variation ratio classifier derived using only the time domain portion of the radar signature. The analysis of the ADAPT base derived using only this portion of the signature, which was presented in Section 3.5, showed that this base was essentially identical to the reference base. Thus, one would expect that the classification algorithms derived on this base would have identical performance. The fact that this has occurred can be verified by comparing Tables 4.2 and 4.3. The mean and standard deviation of each of the classes, for both the learning and independent test cases agree to approximately two places and the same values for the max and min occur for the same cases. Examination of the individual cases confirms that the use of only the time domain for the analysis has had no effect on the classification algorithm. Thus, if the analysis is to be performed using the real and imaginary components without farther non-linear processing it should be performed using only the data from the time domain.

The simplest signature considered in this study from the radar hardware standpoint is that using only the principal polarization in the time domain. Table 4.4 presents the summary of the projection of the 260 learning and 340 independent test cases onto the minimum variation ratio classifier derived using the principal polarization time domain base. Comparison of Table 4.4 and the projection derived from the reference base presented in Table 4.2 shows these two bases to be similar. The similarities can be seen if one realizes that these bases tended to be mirror images and thus the signs associated with each of the classes are reversed. Thus, a positive mean value for the reference base tends to become a negative mean value for the principal polarization time base. Also the standard deviations associated with each class are approximately the same order of magnitude. The mirror imaging also results in an exchange of maximum and minimum values since positive values are now negative and negative values become positive.

The discussion in Section 4.1 showed that the behavior of the minimum variation detection statistic was both non-Gaussian and often showed differences between the 360 case test set and the 1800 case test set. Thus, these algorithms derived on the principal polarization time base were also evaluated against the 1800 independent test cases. The results of this evaluation are presented in the classification trade-off curves

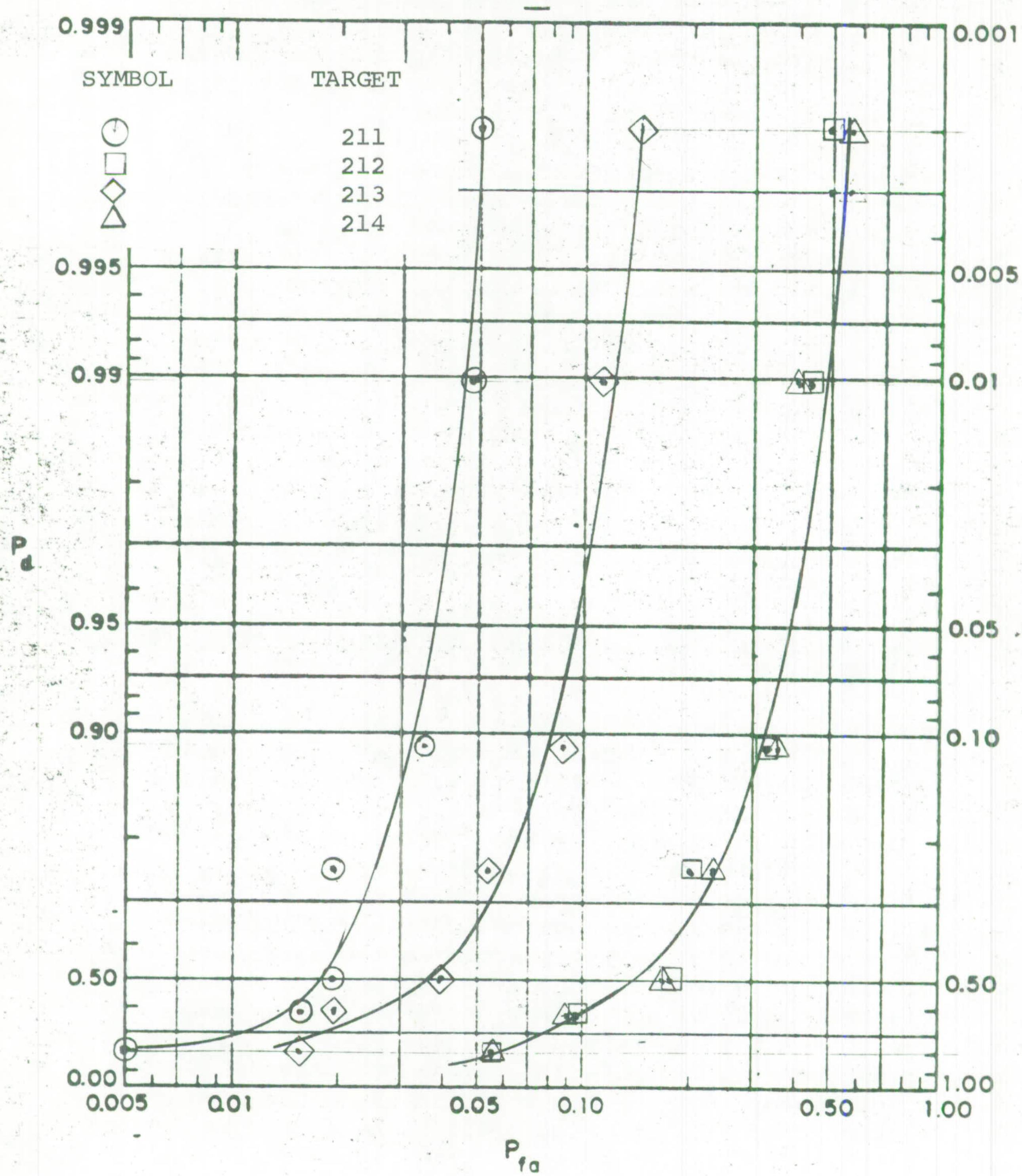
shown in Figure 4.16. Comparison of Figure 4.16 with Figure 4.15 shows the surprising result that the performance has not been degraded by limiting the data to the principal polarization only. The slight improvement in the classification performance seen in this figure is statistically insignificant. One also notes, that in general the relative ease of classifying the targets is the same for the algorithms derived in both the reference base and the principal polarization time domain base.

Assuming no useful information is discarded in the truncation of the Karhunen-Loeve expansion, the above results and the conclusions reached in Section 3 suggest that the performance of any classification algorithm derived using the principal polarization only base would be identical to the performance of the classification algorithms derived using the principal polarization time domain base. This was verified by deriving the minimum variation ratio classifier for the principal polarization only base. Projection of the 260 learning cases and 340 independent test cases on the separation direction for this classifier is summarized in Table 4.5. Comparison of Tables 4.4 and 4.5 verify that these bases produce the same classification performance.

Effect of Variations in Learning Data

The sensitivity of the classification algorithms to additional targets belonging to both the 21X and 101 Class was evaluated by deriving minimum variation ratio classifiers using the Class II only and the 3-Class II target bases. Table 4.6 presents the summary of the projection of the 260 learning cases and 340 independent test cases on the separation direction for the minimum variation ratio classifier derived using the Class II base. Comparison of Table 4.6 with the projection on the separation direction using the reference base presented in Table 4.2 shows that these projections are essentially mirror images of one another. The means of each class are equal to the negative of the means of the corresponding class in the reference base to an accuracy of approximately two decimal places. The standard deviations of the projection of each class also agree for those derived using the reference base to approximately two decimal places. Again, because of the

FIGURE 4.16 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF ALL THE TARGETS ON THE MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE PRINCIPAL POLARIZATION TIME BASE



mirror imaging the cases having maximum values for the Class II only base are exactly those cases having minimum values for the reference base and vice-versa. The values of the max and min cases are also negative of each other. Thus, we conclude that the deletion of the 101 Class from the derivation of the base did not affect the performance of the classification algorithm derived from the base. This is consistent with the results obtained in Section 3 where we saw that the same deletion of Class 214 did not significantly affect the characteristics of the base.

The effect of adding additional members of the 21X Class was evaluated by using the 3-Class II target base to derive the minimum variation ratio classifier. The performance of this classifier is summarized in Table 4.7. Comparison of this table with the performance on the reference base presented in Table 4.2 shows that the effect of using only three members of the Class 21X is very small. The means and standard deviations of the two classes agreed to better than one significant figure. Except for the 212 Class, the cases for which the maximum and minimum occur also agree between the two bases. For the 212 Class, we see differences in the cases which are maximum and minimum. Thus, we conclude that the elimination of 214 Class had a small but slightly greater effect than the elimination of the entire 101 Class from the learning data set.

4.5 Discrimination Based on Linear Classifiers Applied After Square Pre-Processing

The square base was constructed from the same data used for the reference base but after each of the components had been squared. The purpose of this squaring was to create a base in which the mean values for the 101 and the 21X Class would be different. By squaring the difference between the mean of the 101 Class and the value of each of the components this could be achieved. Since the mean of the 101 Class was essentially zero, the simple squaring of the variables was used to accomplish this. After performing the squaring pre-processing and developing the ADAPT optimal base, the Fisher classification direction was derived for this space. The projection of the learning data on the Fisher classification direction is shown in Figure 4.17A, B, and C. Figure 4.17A and B showed the projection of the 21X Class on the Fisher classification direction and Figure 4.17C shows the projection of the 101 Class on this direction. This figure shows the

FIGURE 4.17A - BAR CHART SHOWING THE MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE FISHER CLASSIFICATION DIRECTION FOR THE SQUARE BASE

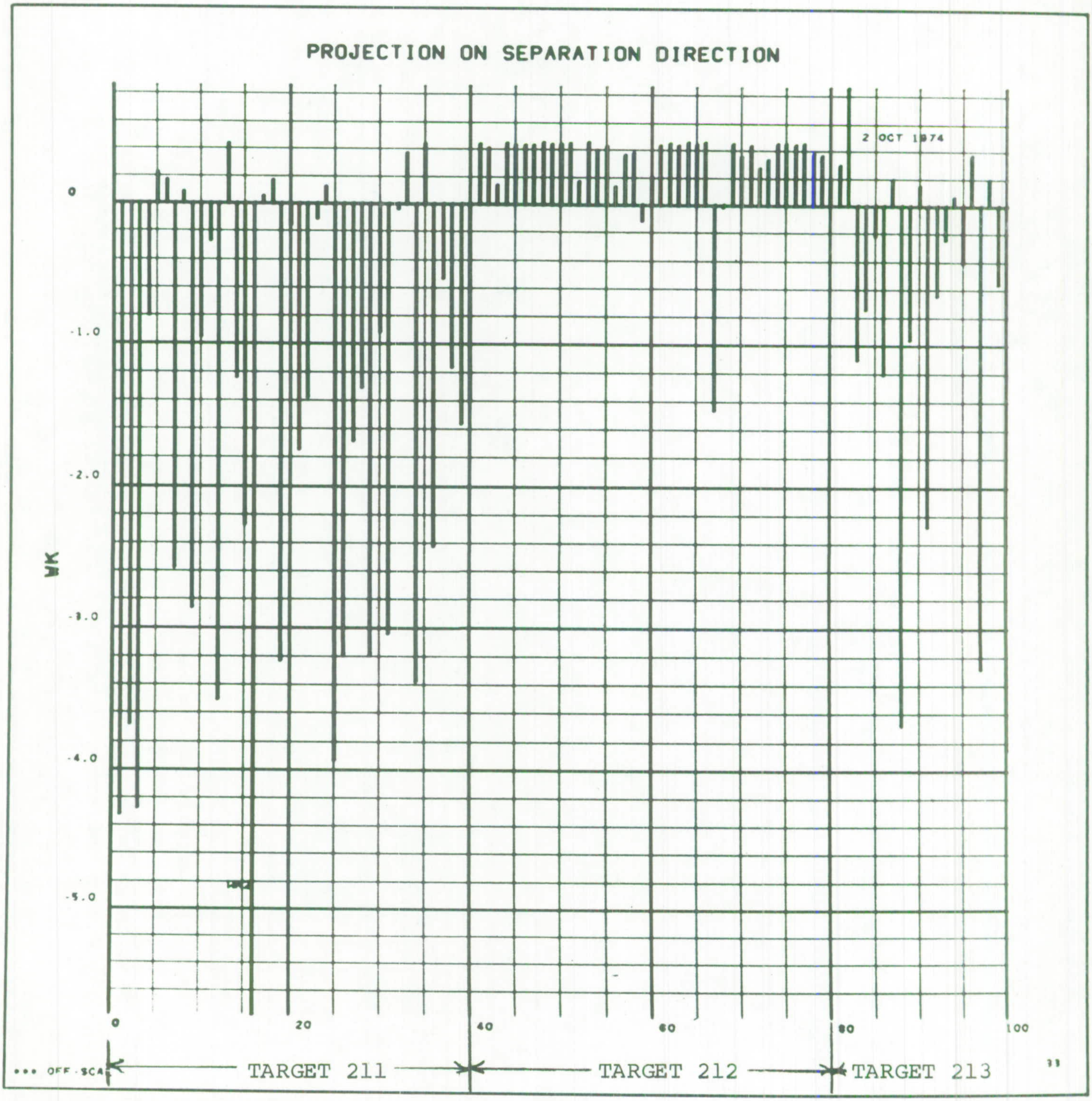


FIGURE 4.17B - BAR CHART SHOWING THE MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE FISHER CLASSIFICATION DIRECTION FOR THE SQUARE BASE

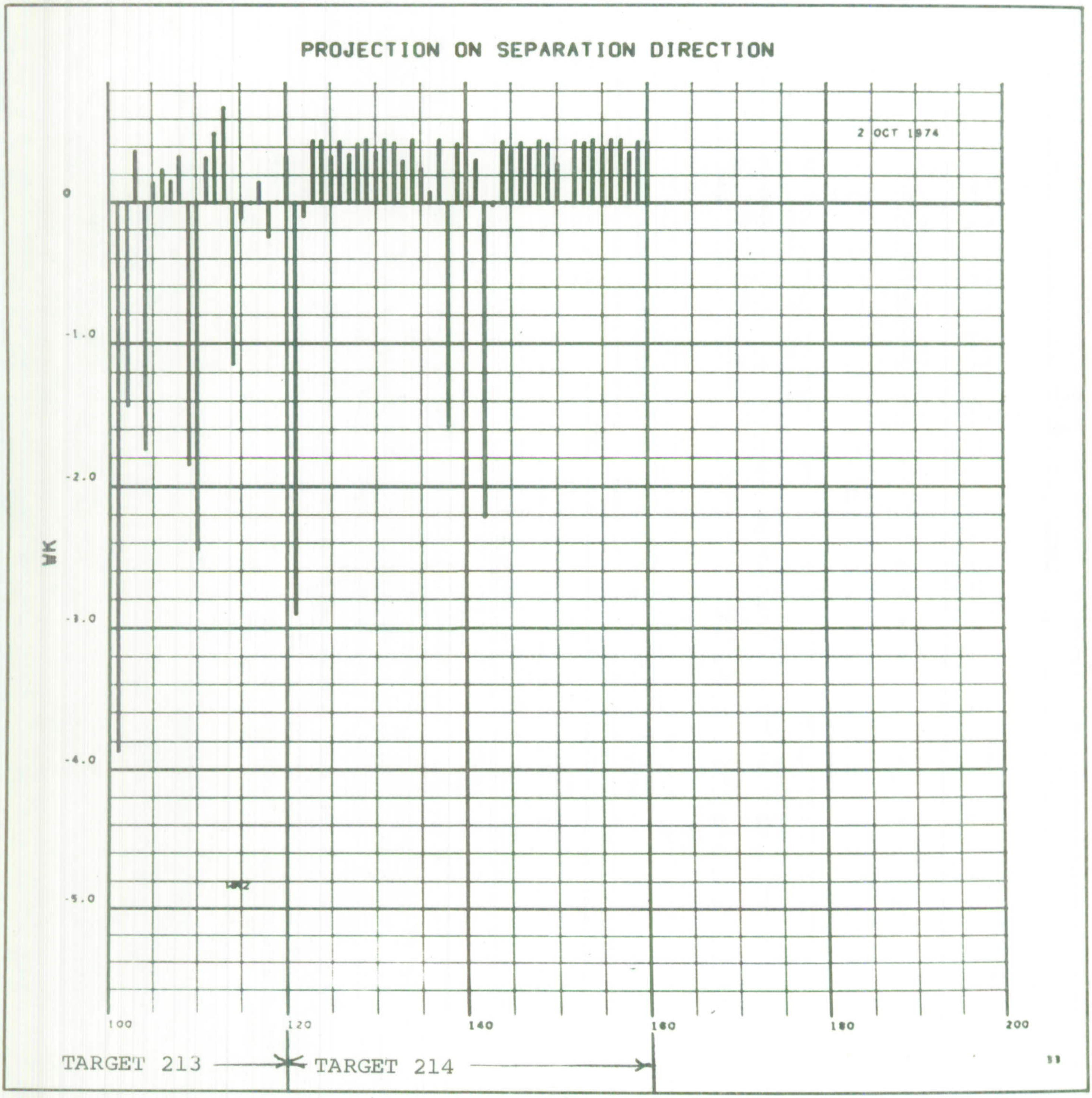
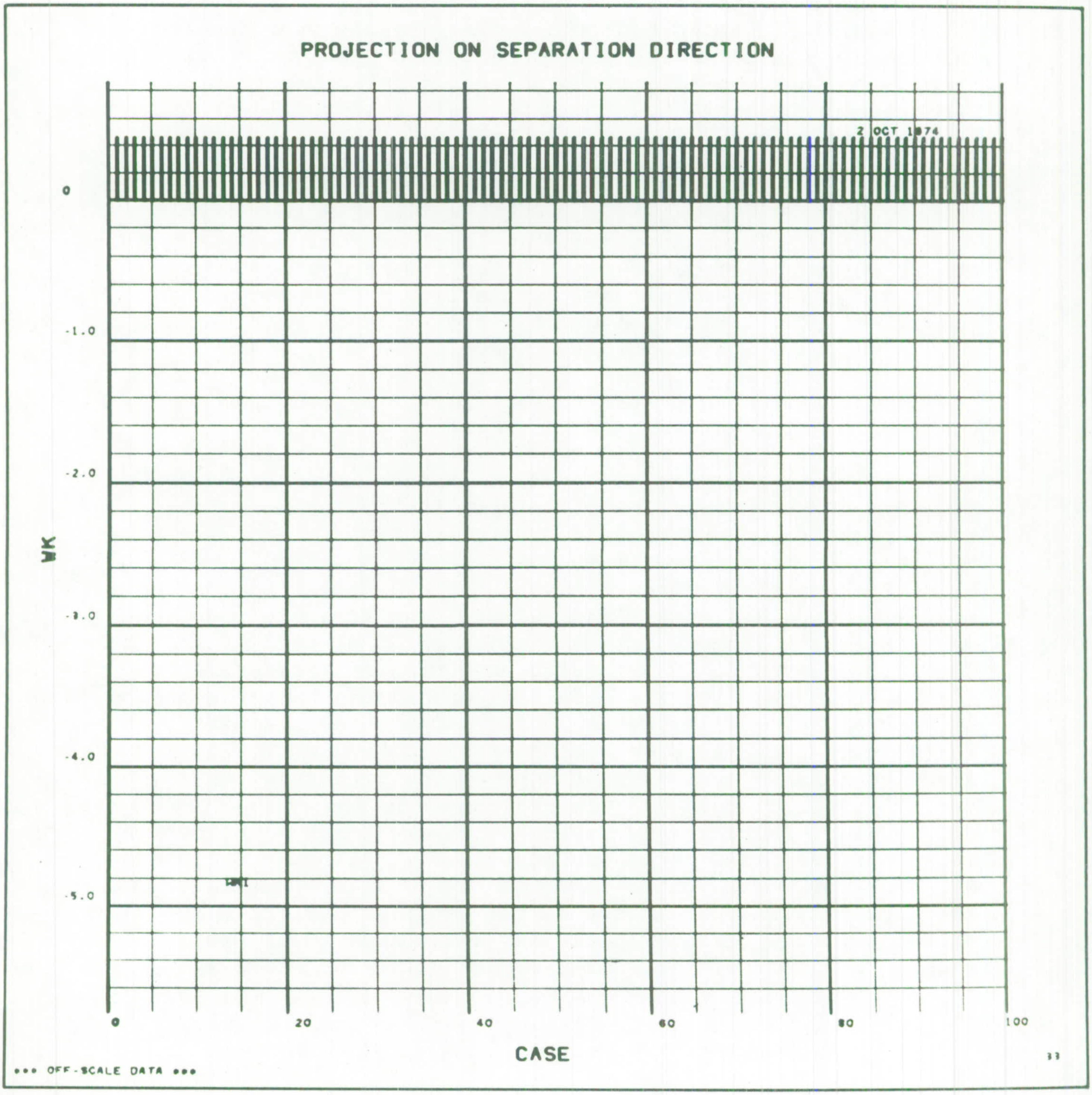


FIGURE 4.17C - BAR CHART SHOWING THE MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE FISHER CLASSIFICATION DIRECTION FOR THE SQUARE BASE FOR TARGET 101



surprising result that although the situation has improved significantly relative to the reference base, the variation of the 21X Class targets still plays a dominant role in the classification. The mean values for the 212 and 214 targets are still very similar to the 101 class. Furthermore, the 101 Class still shows extremely small variation. In fact, detail analysis of numerical results show that the variation is even less than in the reference base.

Figure 4.18 shows the relative importance spectrum for the Fisher classification direction. This spectrum shows that the Fisher classification direction makes use of the higher order optimal directions far more than the lower order optimal directions. The relative importance vector obtained by transforming the Fisher classification direction back to the original data space is presented in Figure 4.19. This figure shows a slight preference for the variables in the time domain although significantly less than was seen in the relative importance vectors using the bases derived without the square pre-processing. This figure also shows that the principal polarization plays a significantly more important role then the opposite polarization.

The performance of the Fisher classification law on the 1800 independent test sets is presented in terms of the classification trade-off curve in Figure 4.20. Figure 4.20 shows that the classification performance for the more difficult targets (i.e. 212 and 214) is approximately the same as was observed with the minimum variation ratio classifier on the reference base. The classification of the easier targets, (i.e. 211 and 213) is significantly improved over that using the minimum variation ratio classifier and the reference base. The results presented in Figure 4.20 may also be compared with the expected performance based on assumption of Gaussian distribution for the detection statistic. Analysis of the Fisher detection statistic shows that if the distributions were Gaussian the detection probabilities for the 211, 212, 213 and 214 Classes should be approximately .27, .41, .30 and .45, respectively, over the entire range of detection probabilities illustrated on Figure 4.20. Although the shape of the experimental curve shown on Figure 4.20 approximates the vertical lines indicated by the Gaussian analysis, the values for the 211 and 213 Classes are more than in order of magnitude in error. Thus, the Fisher detection statistics

FIGURE 4.18 - RELATIVE IMPORTANCE OF EACH OPTIMAL COORDINATE TO FISHER CLASSIFICATION DIRECTION USING THE SQUARE BASE

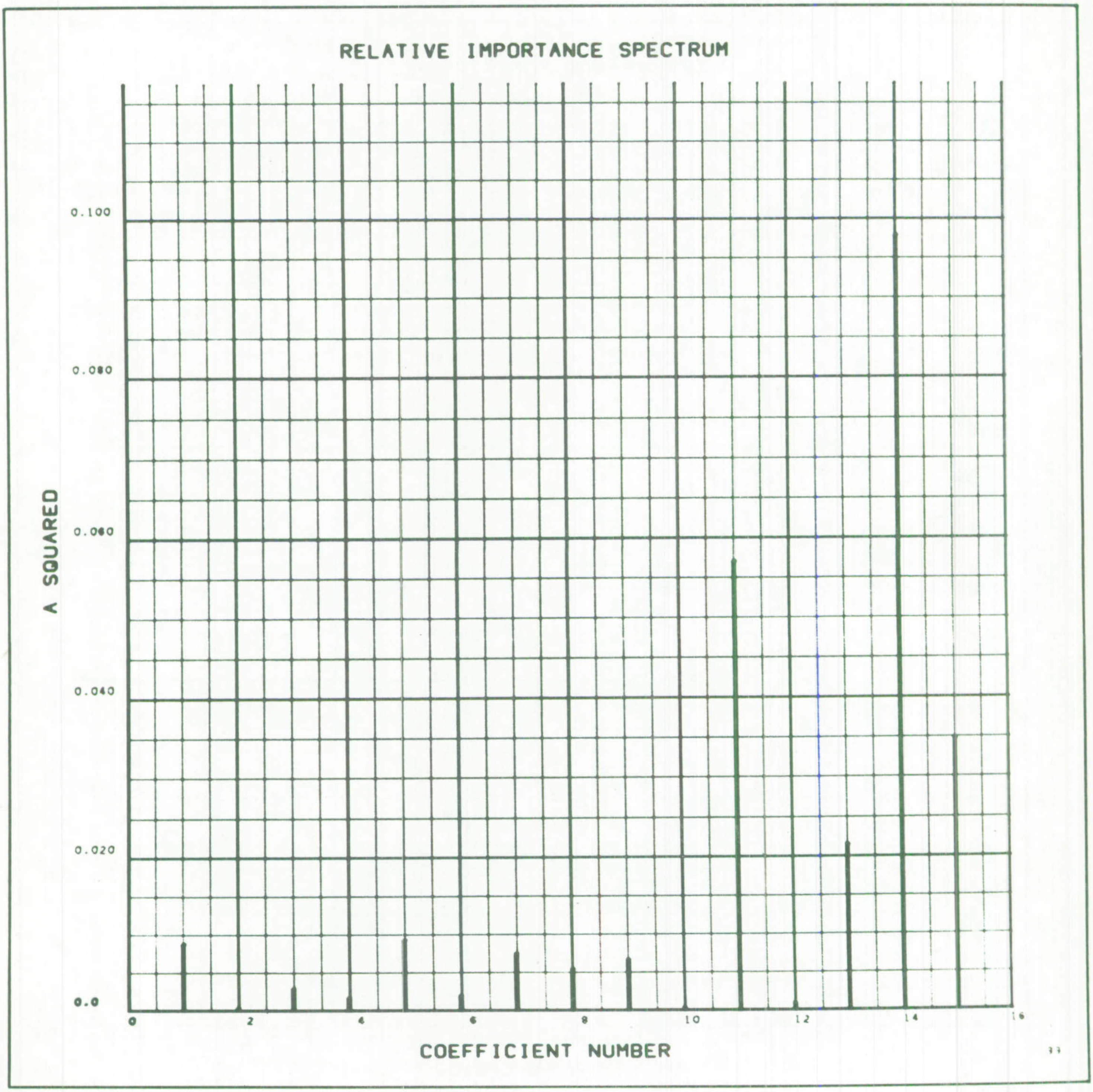


FIGURE 4.19 - RELATIVE IMPORTANCE OF SIGNAL ELEMENT CORRESPONDING TO INDEXING VARIABLE FOR FISHER CLASSIFIER USING SQUARE BASE

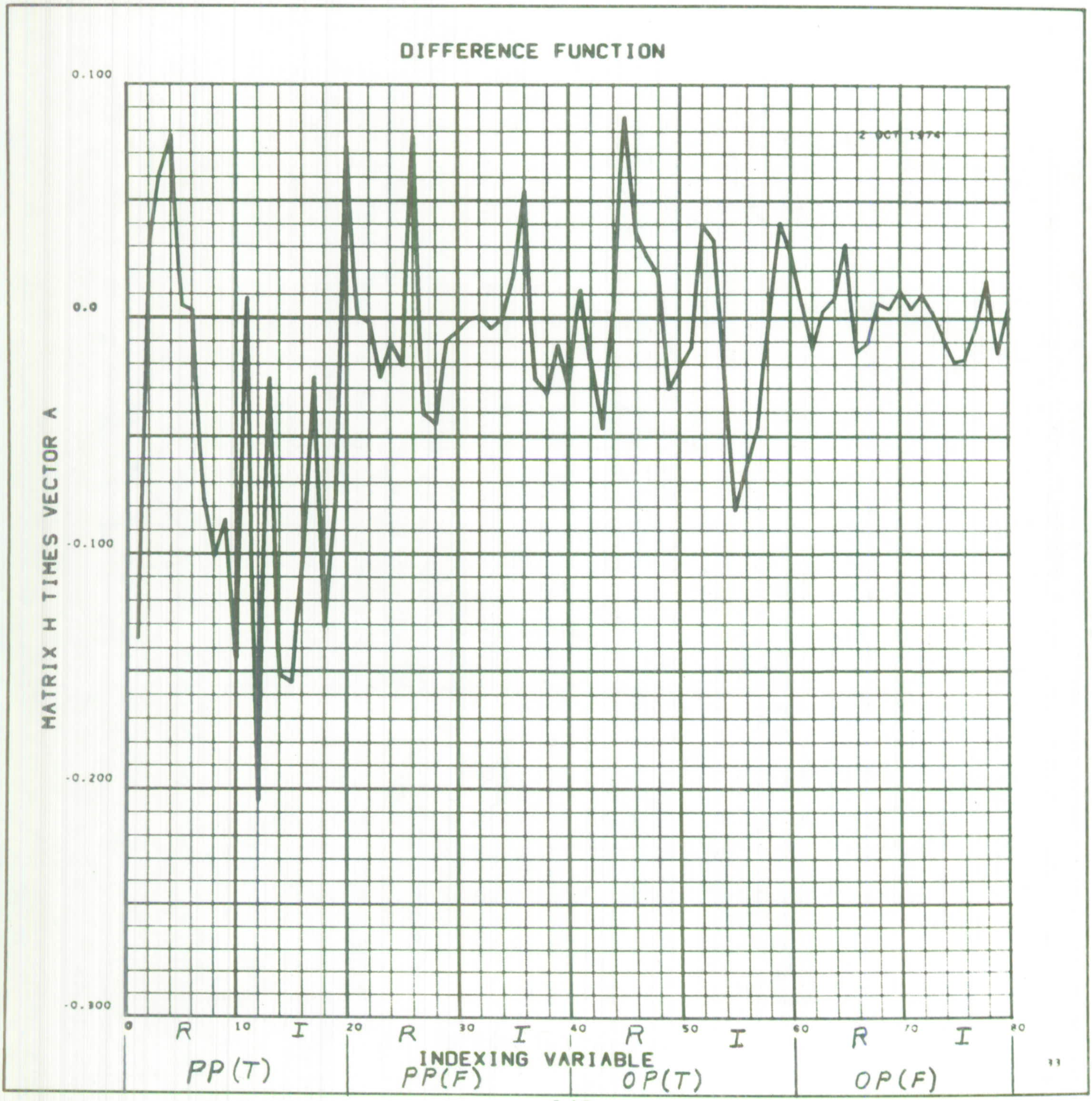
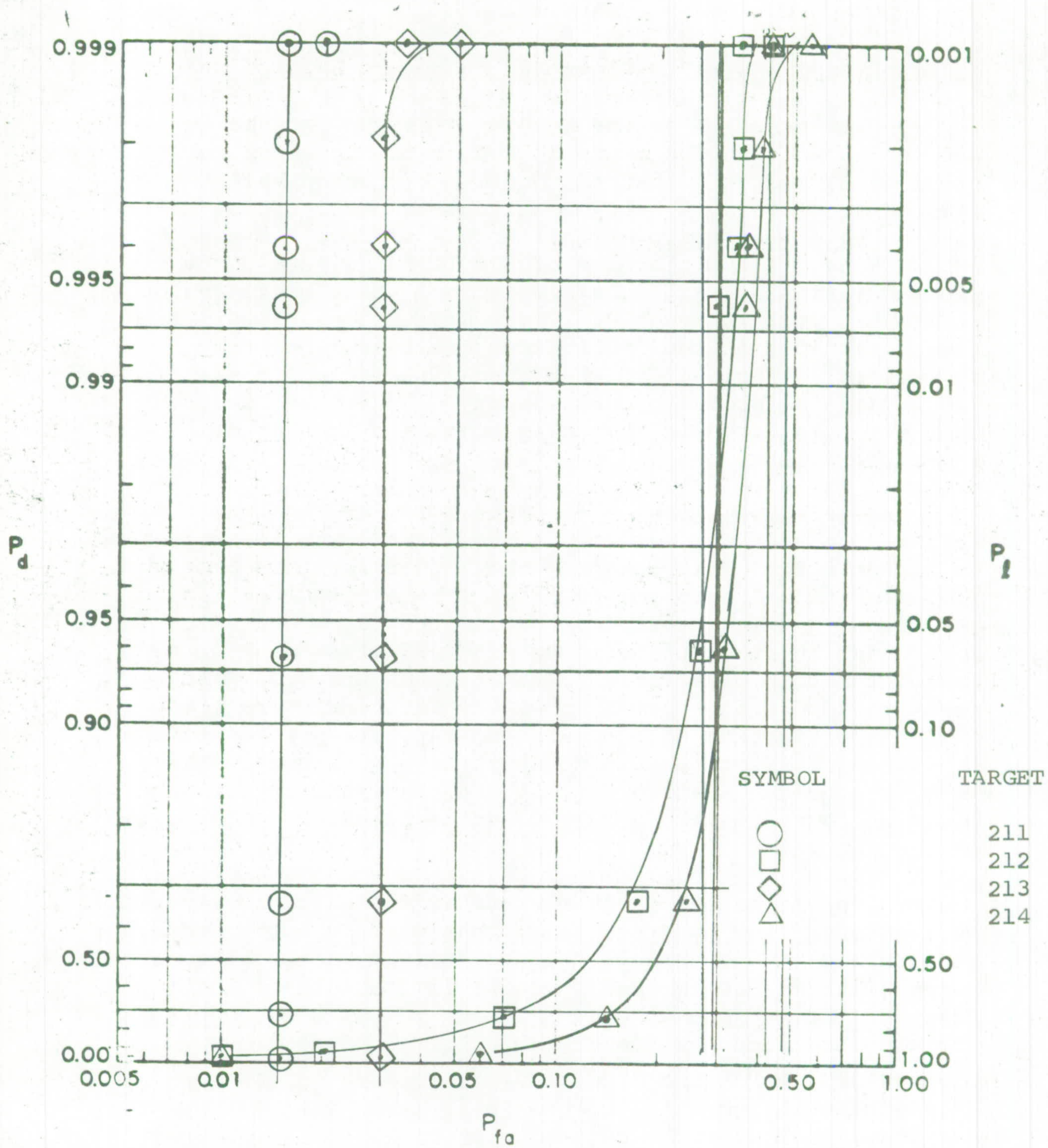


FIGURE 4.20 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF ALL THE TARGETS ON THE FISHER CLASSIFIER DEVELOPED USING THE SQUARE BASE



do not appear to be Gaussian.

Minimum Variation Ratio Classifer on Square Base

The examination of the scatter plots and of the projections of the Fisher detection statistic shown in Figure 4.17 suggests that the minimum variation ratio classifier might work well on the square base. Thus, the minimum variation ratio classifier was derived for the square base. The projection of the learning cases on the minimum variation ratio classification direction is shown in Figures 4.21A through C. These projections as well as the projections of the 360 independent test cases are summarized in Table 4.8. When comparing Table 4.8 and Figure 4.21 an additive constant of -0.0828 must be added to the values given in the table to obtain the values shown in Figure 4.21.

The relative importance spectrum for the minimum variation classifier derived on the square base is presented in Figure 4.22. This spectrum shows that for this case the classification direction is primarily in the direction of the second optimal coordinate but many other coordinates make significant contributions. The relative importance vector for this classifier is shown in Figure 4.23. This relative importance vector shows the interesting result that this classification is dominated by the time portion of the opposite polarization. In contrast to the reference base, we must conclude that for the square base, the performance would be degraded if one only used the principal polarization of the return.

The performance of the minimum variation ratio classifier derived using the square base is considerably better than the performance of the same classifier derived using the reference base. This is easily seen by comparing the classification trade-off curves presented in Figure 4.24 with those presented for the reference base in Figure 4.15. In fact, the performance on the 211 and 212 target is better than can be evaluated using the 1800 case test set. However, one may state that over the detection probability ranges shown on Figure 4.24, the 211 and 213 targets most likely have false alarm rates of the order of half a percent or less.

FIGURE 4.21A- BAR CHART SHOWING THE MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION DEVELOPED USING THE SQUARE BASE

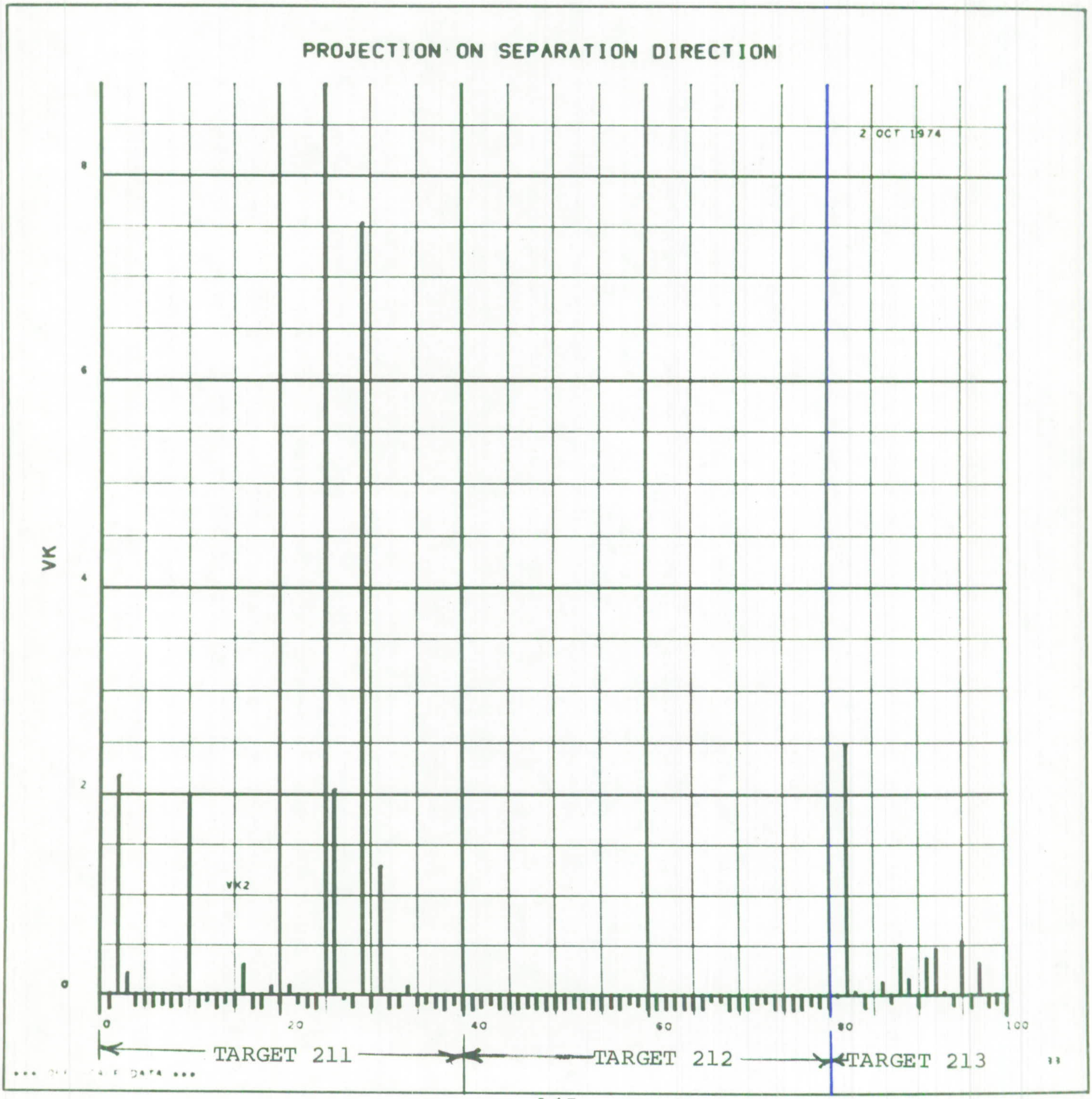


FIGURE 4.21B - BAR CHART SHOWING THE MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION DEVELOPED USING THE SQUARE BASE

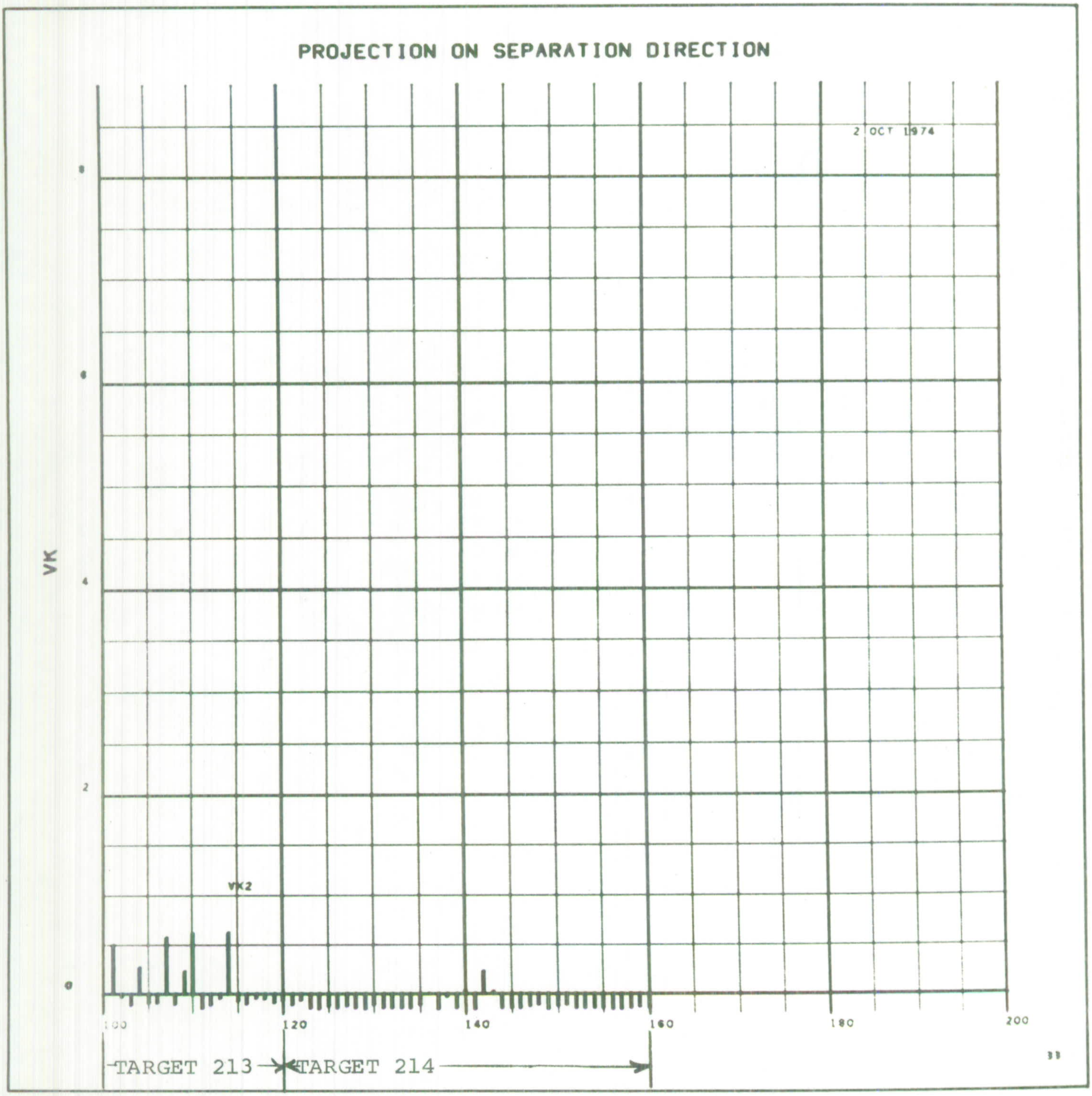


FIGURE 4.21C - BAR CHART SHOWING THE MAGNITUDE OF THE PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION DEVELOPED USING THE SQUARE BASE FOR TARGET 101

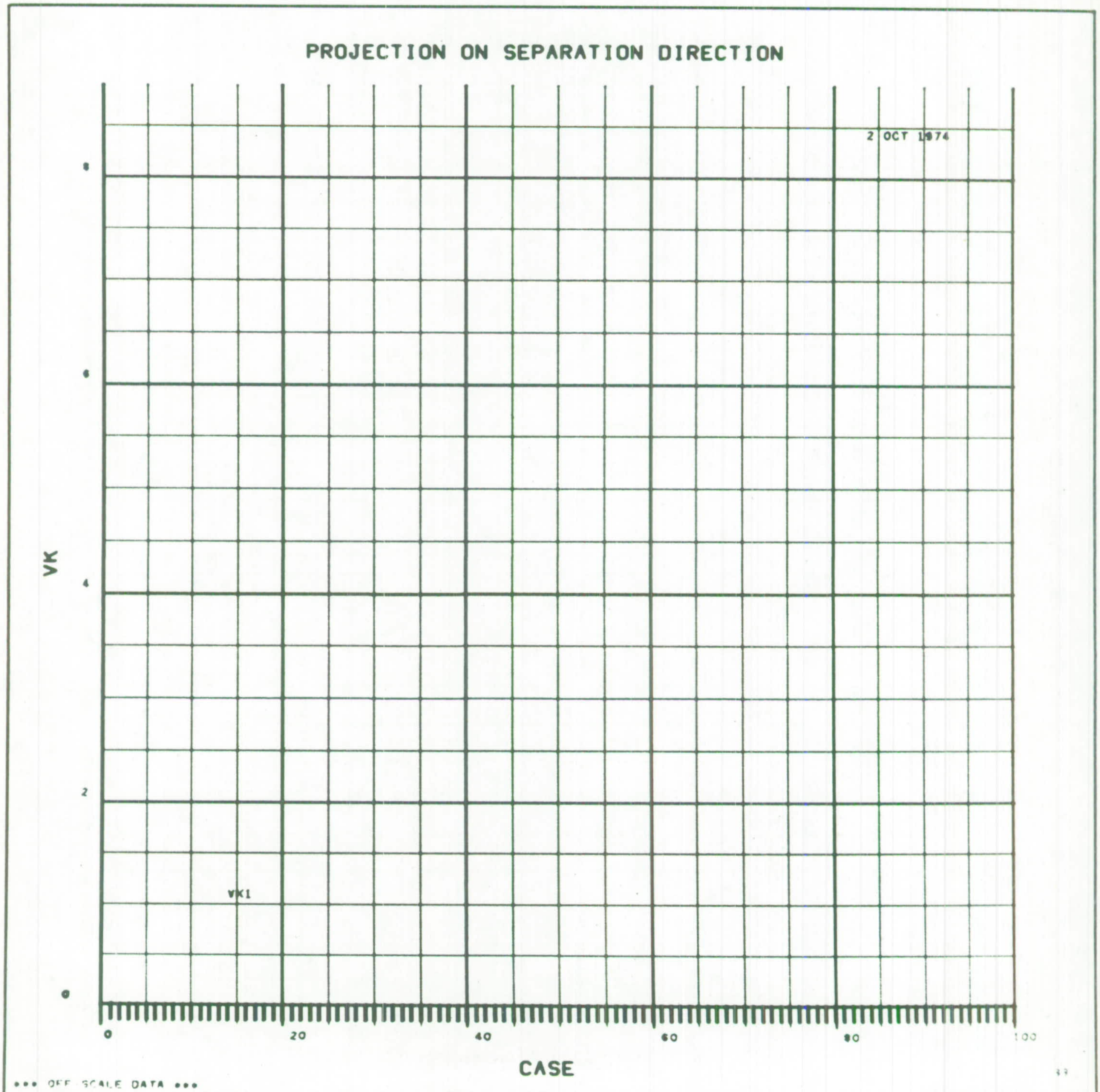


FIGURE 4.22 - RELATIVE IMPORTANCE OF OPTIMAL COORDINATES TO MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE SQUARE BASE

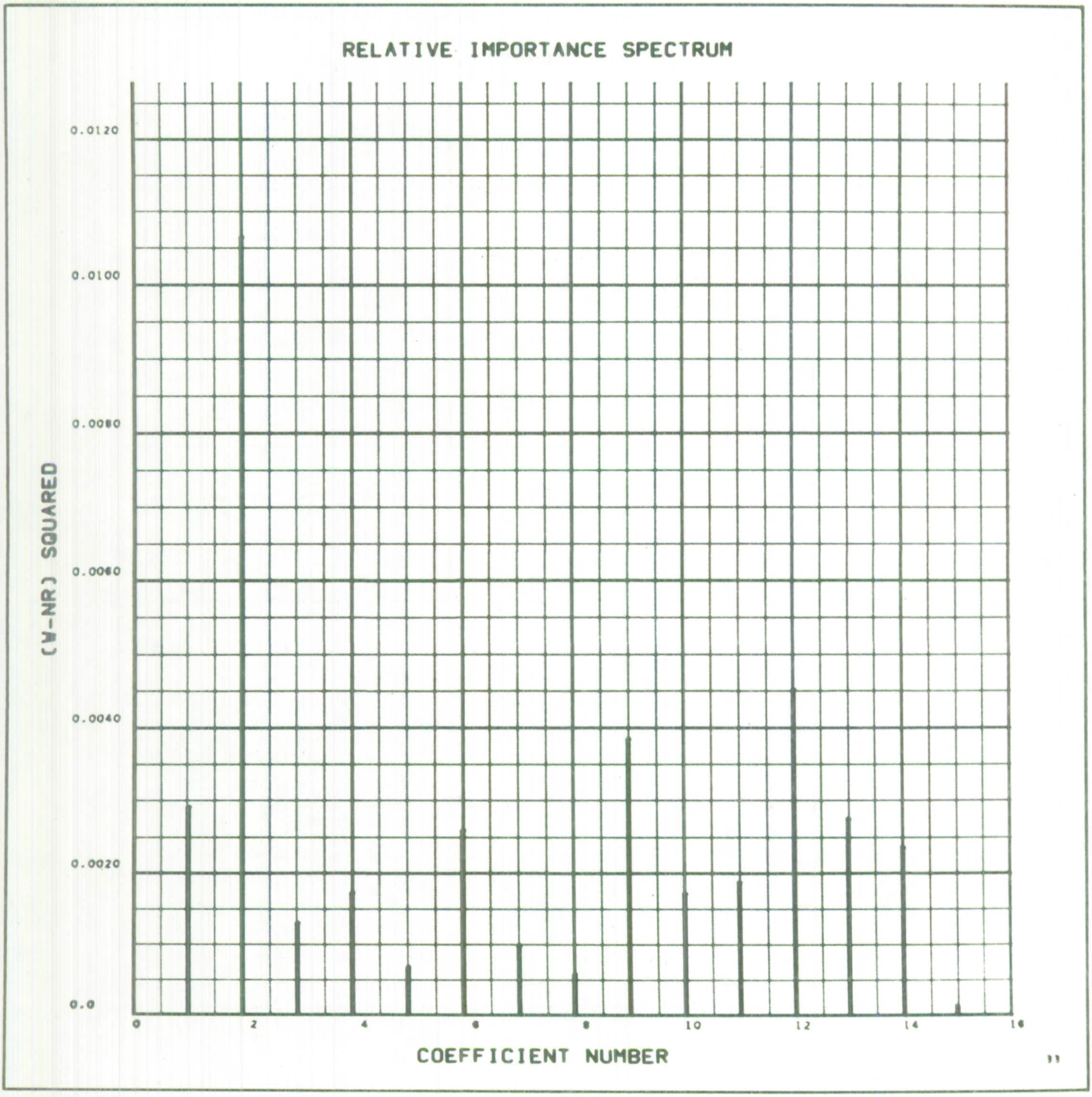


FIGURE 4.23 - RELATIVE IMPORTANCE OF SIGNAL ELEMENT CORRESPONDING TO INDEXING VARIABLE FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE SQUARE BASE

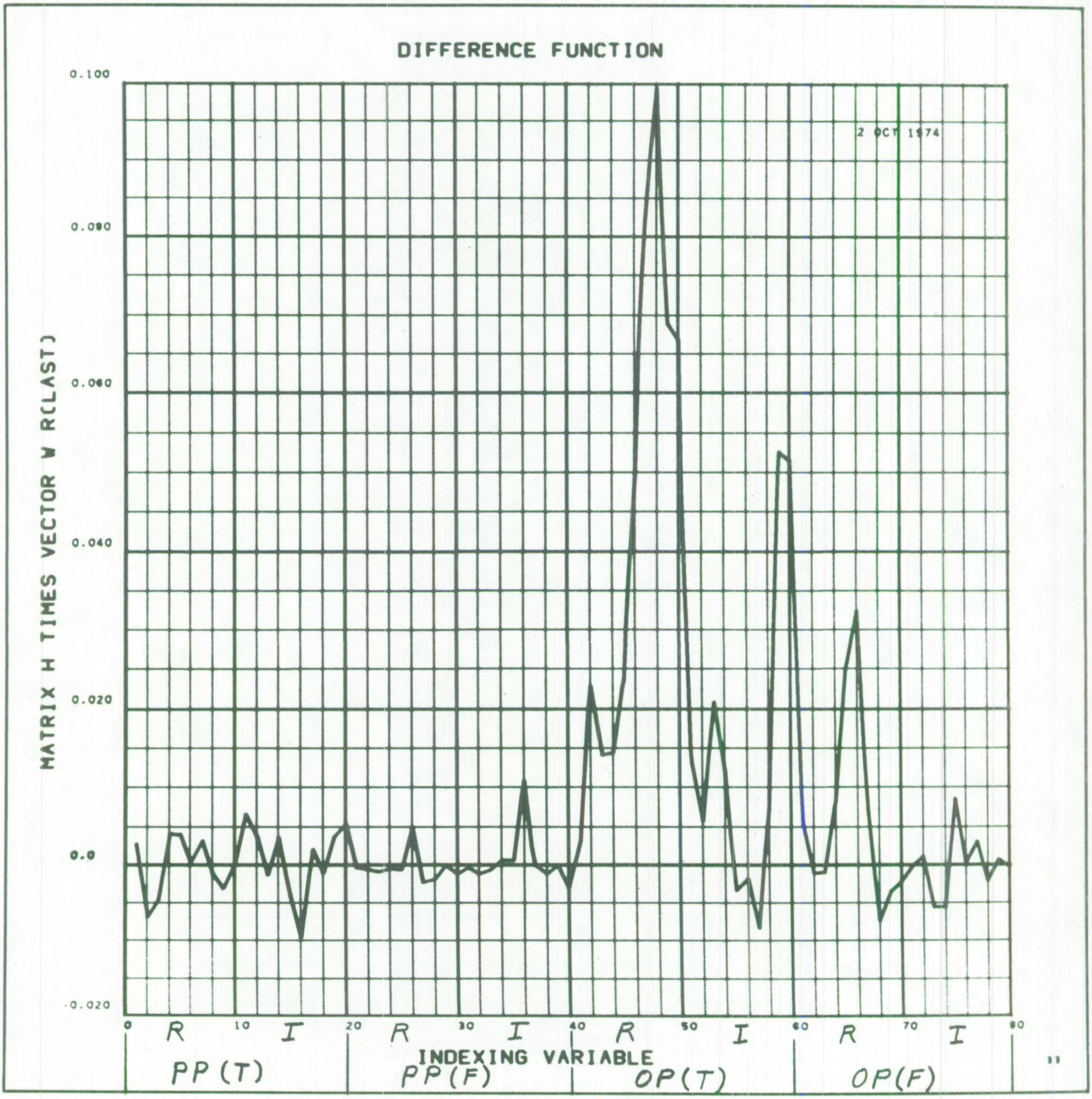
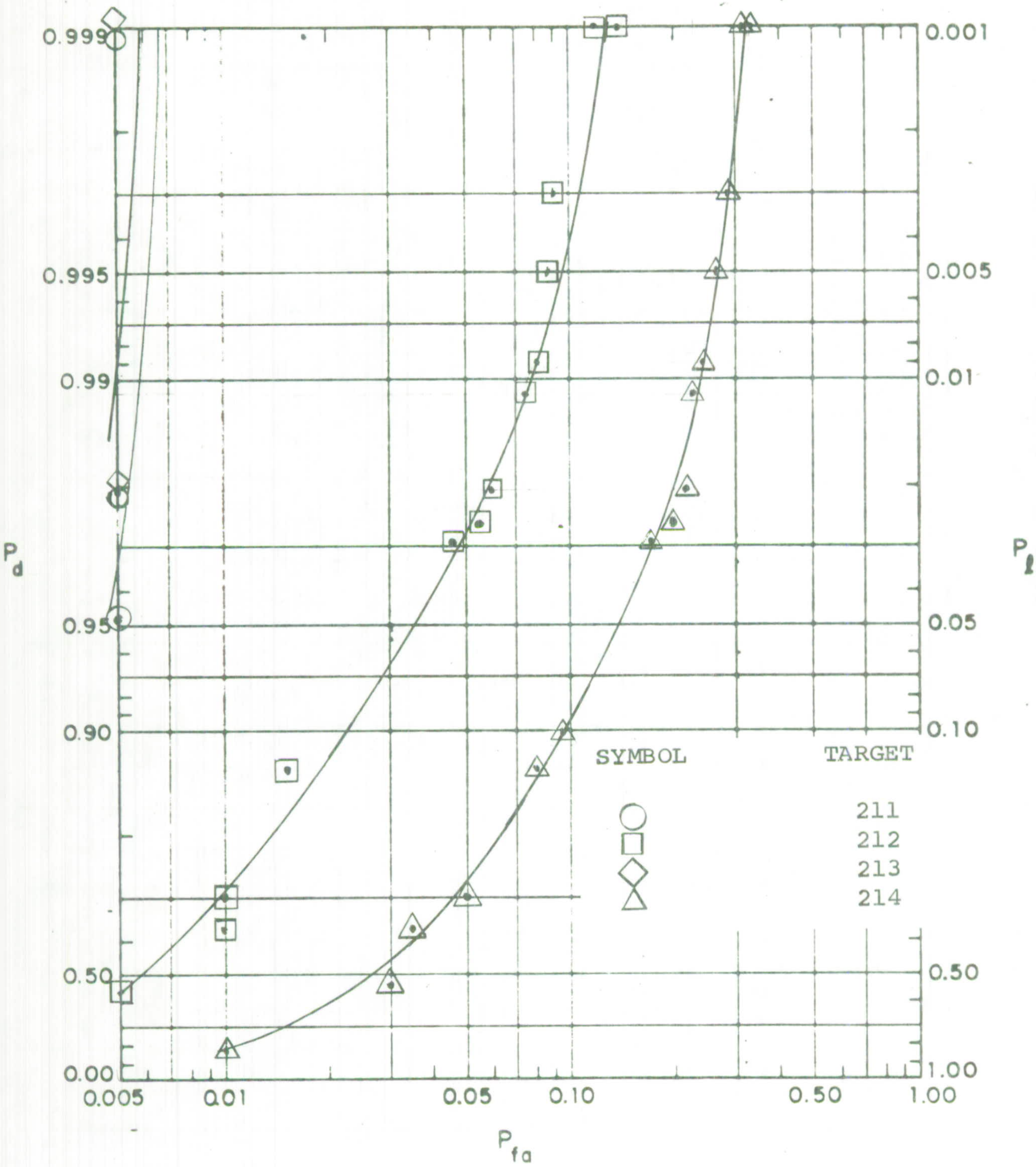


FIGURE 4.24 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF ALL THE TARGETS ON THE MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE SQUARE BASE

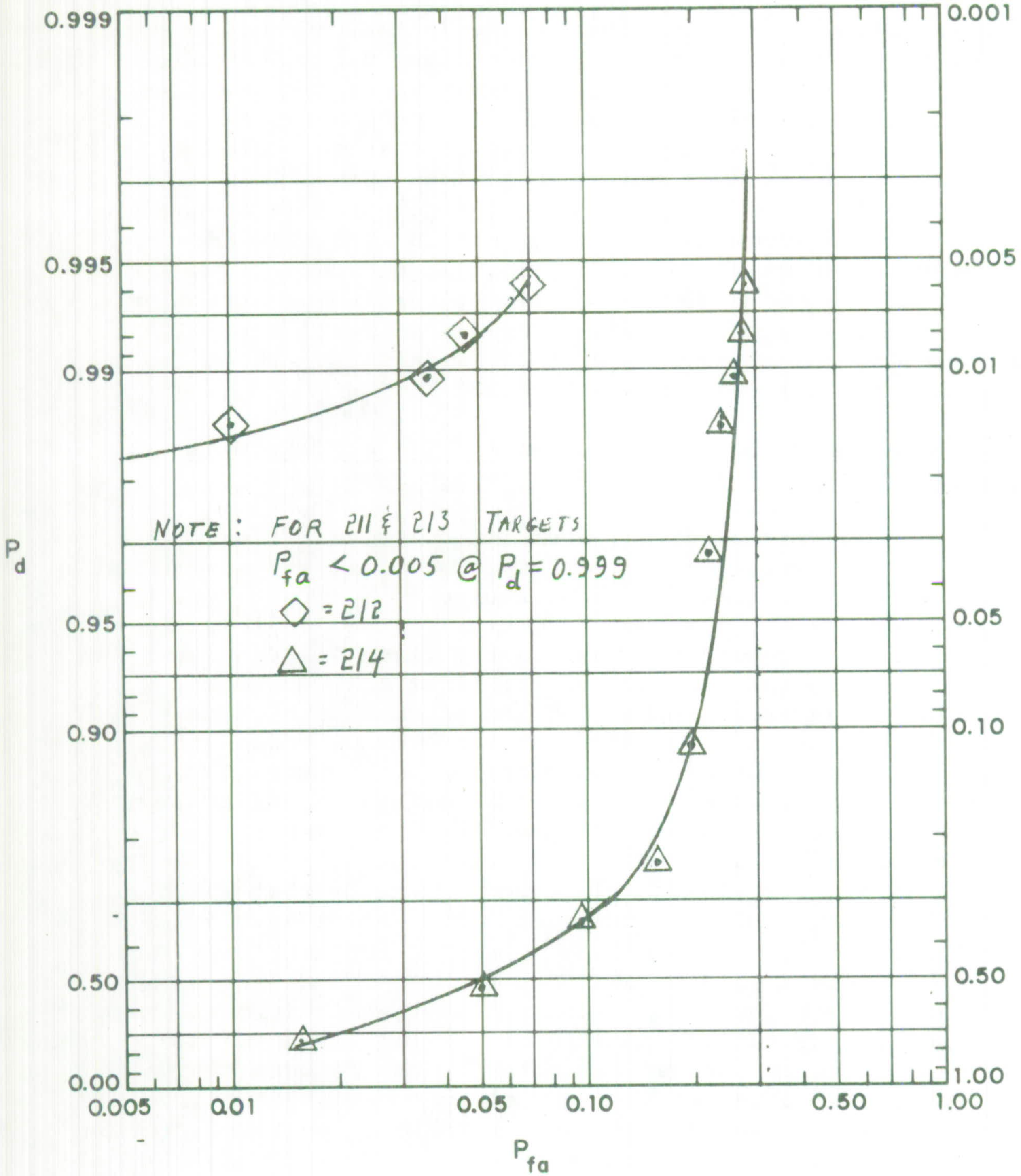


The analysis of the square base and the base obtained by subtracting the mean of Class I prior to squaring presented in Section III showed that no significant difference between these two pre-processings were visible in terms of the characteristics of the bases derived. To verify that this similarity was still present when one used these bases to derive classification laws, the minimum variation ratio classifier was derived for the zero mean square base. The projection of the 260 learning and 340 independent test cases on the minimum variation ratio separation direction for this base is summarized in Table 4.9. Comparison of Table 4.9 and Table 4.8 shows that the mean values of the classes are in agreement to two significant figures and the standard deviations to more than four significant figures. Again, the max and min cases are identical and have similar values. Thus, we conclude that the classification results obtained on the square base will be identical to those obtained on the zero mean square base.

<u>Classification Using First Optimal ADAPT Coefficient As Detection Statistic for Zero Mean Square Base</u>

Since the first optimal coefficient proved an effective detection statistic for the reference base, it was also tried as a detection statistic for the square base. Table 4.10 presents the summary of the projection of the 260 learning cases and 340 independent cases on the first ADAPT optimal coordinate. The performance of the first ADAPT optimal direction classifier on the 1800 independent test cases is presented in the classification performance trade-off curve shown in Figure 4.25. Comparison of Figure 4.25 with Figure 4.24 shows that this detection statistic has the same or slightly better performance then the minimum variation ratio classifier. It should be pointed out that the performance where the 211 and 213 targets exceeds that which can be measured using 1800 case test set. The fact that no false alarms occur even when all 1,000-101 test cases are correctly identified suggests that the use of the first ADAPT coefficient as a detection statistic may be significantly better than the minimum variation ratio classifier for these two targets. This conjecture is further supported by the fact that this result was also obtained when the minimum variation ratio classifier was compared with the first ADAPT optimal direction classifier on the reference base.

FIGURE 4.25 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF ALL THE TARGETS ON THE FIRST ADAPT OPTIMAL DIRECTION CLASSIFIER DEVELOPED USING THE SQUARE BASE



Prognosis For Use of Incoherent Signature

The major difference between the data used for the Task 1 and the Task 2 studies was that the Task 1 data was incoherent data and the Task 2 data was coherent. One important task which remains to be performed is to repeat a portion of the classification analysis of the Class II study using only the amplitude of the return to assess the effect of this difference on the performance. This analysis should be carried out parallel to that performed using the square variable base. To perform this, one would add the square of the real and imaginary portions of the return, thus, reducing the signature to one half the number of points. This would be equivalent to processing the square of the incoherent amplitude. These results could be compared directly with the results obtained using the square base. The differences in the performance observed should be attributable to the loss of information in using an incoherent rather than a coherent signature. Although this analysis has not been performed, review of the data vectors, optimal functions and relative importance vectors associated with the square base, suggest that there will be a significant loss of information in utilizing the incoherent signature.

In Section 3, Figures 3.45 and 3.46 gave sample data histories using the squared variables. The average of all 260 learning cases used to construct the square base was given in Figure 3.47. The corresponding incoherent signature may be constructed from the time domain portion of the signature by adding the corresponding indexing variables in the set determined by indexing variables 1 through 10, and 41 through 50 to the set defined by indexing variables 11 through 20, and 51 through 60. If two values to be added both have the same effect, then the sum will be as effective as the individual value, however, if one has a positive effect and the other a negative effect, the information conveyed by this difference will be degraded.

We may examine the optimum functions for the square base which are given in Figures 3.49 and 50 and 3.53 to see the role that the variables play in defining the variation associated with the data set. Examination of the first two optimal functions shows a strong spike associated with Indexing Variable 5 and the variable which would be added to

this variable (i.e. Indexing Variable 15). Both the typical histories and the average input vectors as well as the first two and fifteenth optimal functions show that these two variables act in opposite directions. Examination of the set of optimal functions shows this to be true for the majority of the optimum functions. The summation of these two indexing variables will result in the loss of a significant portion of the organized variation from the data set. This suggests that if one were to perform the analysis on the incoherent signature there would be information lost if only the incoherent signature were used. This was observed in the studies presented in Task 1.

The relative importance vectors presented in Figures 4.19 and 4.23 show very little similarity in behavior between the frequency domain and the time domain. Thus, one suspects that a considerable portion of the information being used for both the Fisher classification and the minimum variation ratio classifier will be lost to the incoherent processing. This suggests that these classifiers will behave very differently with this data and will probably have a somewhat worse performance.

4.6 Sequential Classifiers

In general, the performance of good classification schemes can be enhanced by repeated application of the scheme on independent test samples or the repeated application of independent classification schemes on the same test sample. We shall define this procedure as the multiple step classification. The classification algorithms developed in this study have performance which make them suitable for this type of application.

In general, if one sequentially applies a second classification scheme for reaching the same decision to the cases assigned to one class in the initial classification, the

combined detection probability given by $P_{D\Sigma}$ may be calculated under the assumption that the individual detection probabilities are independent by taking $P_{D\Sigma}$ as the product of the two detection probabilities or:

$$P_{D\Sigma} = P_{D1} * P_{D2} \tag{4.1}$$

Similarly, under the same assumption the combined false alarm rate $P_{FA\Sigma}$ is given by:

$$P_{FA\Sigma} = P_{FA1} * P_{FA2} \tag{4.2}$$

In this section, we shall consider two different approaches to using linear multiple step classification schemes. The first will be designating the successive signal multi-step procedure. In this procedure, if a target is identified as a member of the 21X Class, it is rejected, but every target that is identified as a member of the 101 Class is re-examined at a later time. If after a sufficient length of time, the two signatures are independent, one may use equations 4.1 and 4.2 to calculate the performance. The performance of the second application is assumed to be given by the same ROC curve as the original application of the algorithm. Thus, if Equations 4.1 and 4.2 above may be used to compute the new expected classification trade-off curve, Figures 4.26 and 4.27 present the successive signal two step performance estimate for the first ADAPT optimal direction classifier applied to the square and reference bases, respectively. Both of these figures show considerable improvement over the corresponding single step algorithms. For example, comparison of Figures 4.15 and 4.26 shows that if one desires a detection probability of .99 the two step successive signal algorithm will yield a false alarm rate of less than .3 for all the fragment types considered. The single step algorithm gives a false alarm rate of slightly over 50%. Similarly, for the first ADAPT optimal direction classifier on the square base, and a detection probability of 0.99, the false alarm rate is reduced from approximately 0.3 to 0.1 when one uses the successive signal multi-step procedure instead of the single step procedure.

FIGURE 4.26 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF ALL TARGETS ON THE TRACKING MULTI-STEP CLASSIFIER DEVELOPED FROM THE MINIMUM VARIATION RATIO CLASSIFIER ON THE REFERENCE BASE

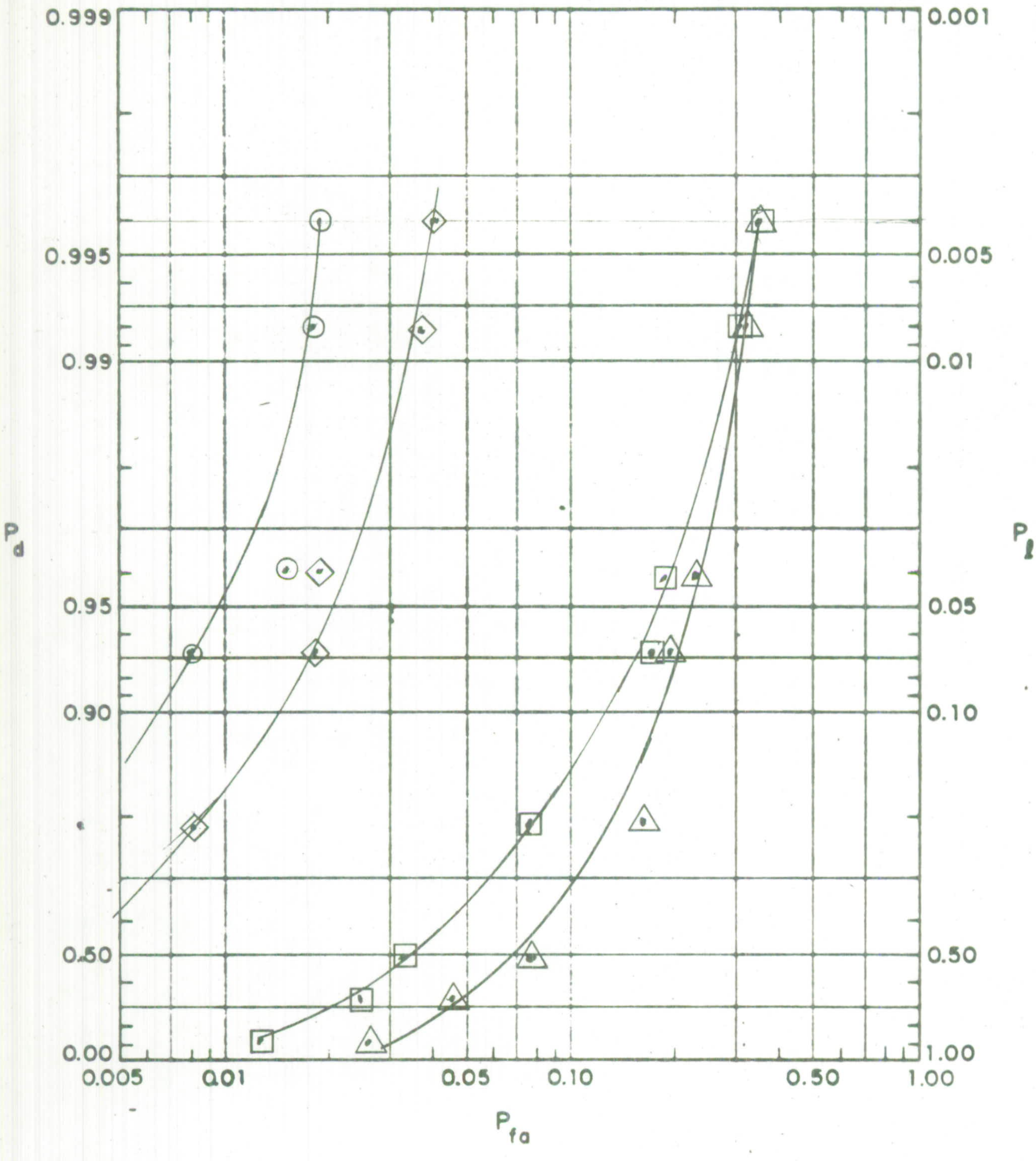
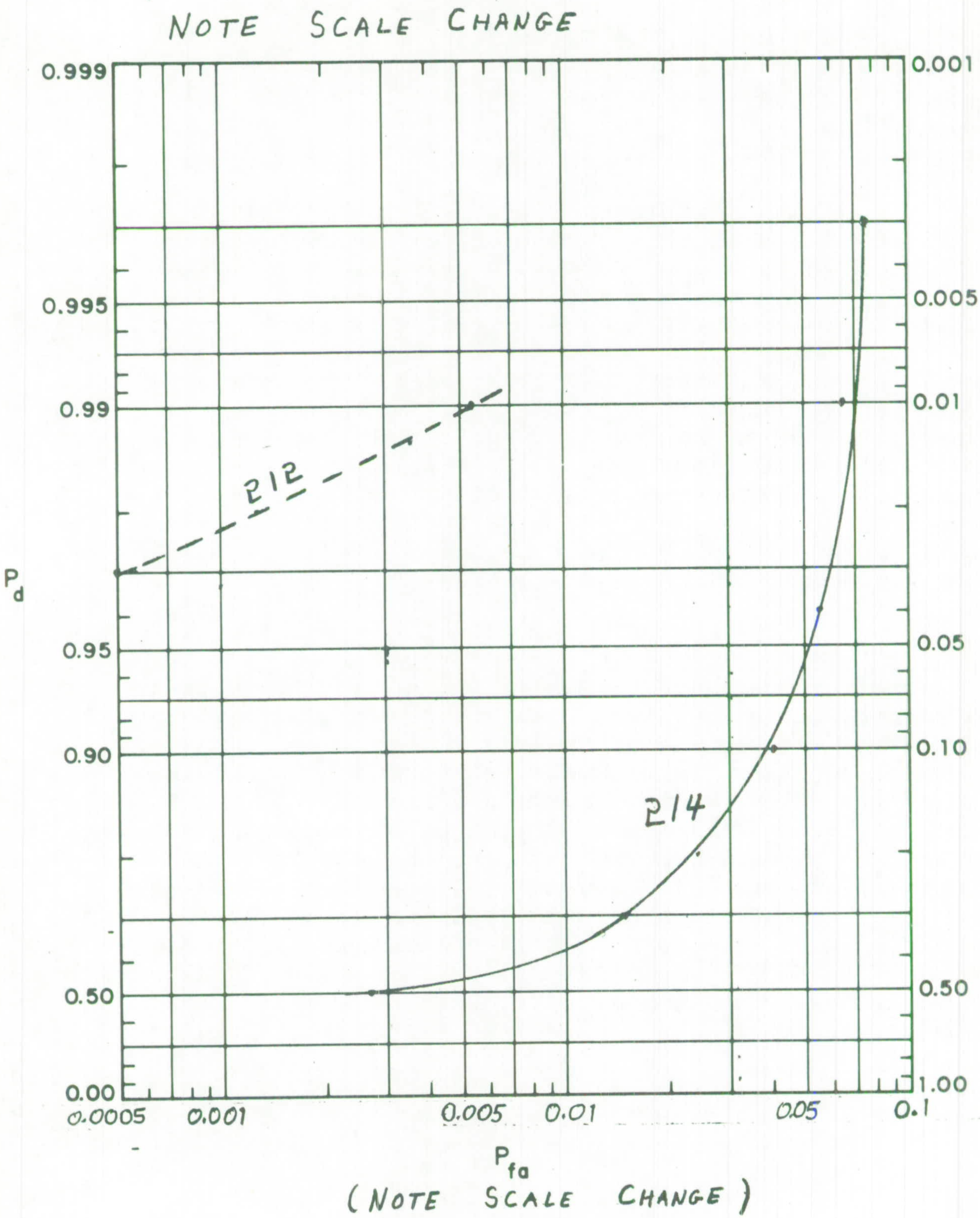


FIGURE 4.27 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF ALL THE TARGETS ON THE TRACKING MULTI-STEP CLASSIFIER DEVELOPED FROM THE FIRST ADAPT OPTIMAL DIRECTION CLASSIFIER ON THE SQUARE BASE



The penalty associated with the application of the successive signal multi-step procedure is the requirement to re-examine those Class II targets which failed the first step procedure. The uncertainty in developing this procedure is that of defining the correlation time between to successive signals. Physical consideration suggest that this should be of the order of the tumbling frequency of the targets.

The majority disadvantage of the successive signal multi-step classification procedure, can be overcome if instead of re-applying the same classification scheme one applies an independent classification scheme to the same data set. In this case, one can achieve the benefits of the two step procedure without the penalty of collecting a second data sample from the target. This scheme could be applied using any combination of the linear classifiers which are discussed here. One approach is to use the second ADAPT optimal direction as the second linear classifier. The analysis of the scatter plots presented earlier showed that all of the first fifteen ADAPT optimal directions could be expected to give reasonably good classification performance.

Another source of algorithms which could be applied as second step algorithm and has the potential for yielding algorithms with equal or even better performance than the original algorithm is to re-derive the algorithm utilizing only those cases which have been misclassified by the first step algorithm. There are two options for accomplishing this when one uses the ADAPT approach to data analysis. The first and simplest is to use the same base as was used for the original algorithm but simply re-derive the algorithm for the new data set. The second approach is to use the new data set to develop a new base as well as a new classification algorithm. Both of these approaches have been performed for the reference base.

Figure 4.28 shows the projections of the learning data onto the minimum variation ratio classification direction developed using the 16 Class 21X cases most like the 101 class in the training set of data. These cases were selected by application of the first ADAPT optimal direction classifier followed by the application of the second ADAPT optimal direction classifier. Note, that this is essentially an examination of the ADAPT scatter plot. This figure shows that a classification algorithm has been developed which can separate all but one of the 16 cases that would have been missed in the scatter plot analysis. Figure 4.29 shows the relative importance spectrum for this classifier. Comparing this with the relative importance spectrum shown in Figure 4.14, we see that the first optimal function no longer plays a significant roll. This is expected since the classification potential of these first two optimal functions was already utilized as part of the scatter plot analysis. Figure 4.30 presents the relative importance vector for the new classification algorithm. It also shows different features than shown in Figure 4.14 and verifies that we have obtained a different classification algorithm by this procedure.

If we assume the classifier defined by Figure 4.30 is applied as a second step after the minimum variation ratio classifier developed on the reference base the performance for this two step classifier is shown in the classification trade-off curve presented in Figure 4.31. The performance is better than that obtained using the minimum variation ratio classifier developed on the reference base. The performance is not as good as that shown in Figure 4.26 for the tracking two step classifier using the minimum variation classifier developed on the reference base.

The same cases which were used to develop the classification algorithm presented in Figures 4.28 through 4.30 were used to develop a complete new ADAPT base. This base is the low variation base which was described in Section 3.3. This base was also used to develop independent classification algorithms. The performance of this algorithm on the training set of data is illustrated by the bar charts presented in Figure 4.32. These figures show that using the new base two of the sixteen cases were missed. Since the development of the low variation base is a considerably greater task then the simple development of the algorithm on the original base and since the performance is similar or poorer this

FIGURE 4.28A - BAR CHART SHOWING THE MAGNITUDE OF PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE LOW VARIATION CASES ON THE REFERENCE BASE - 21X TARGETS

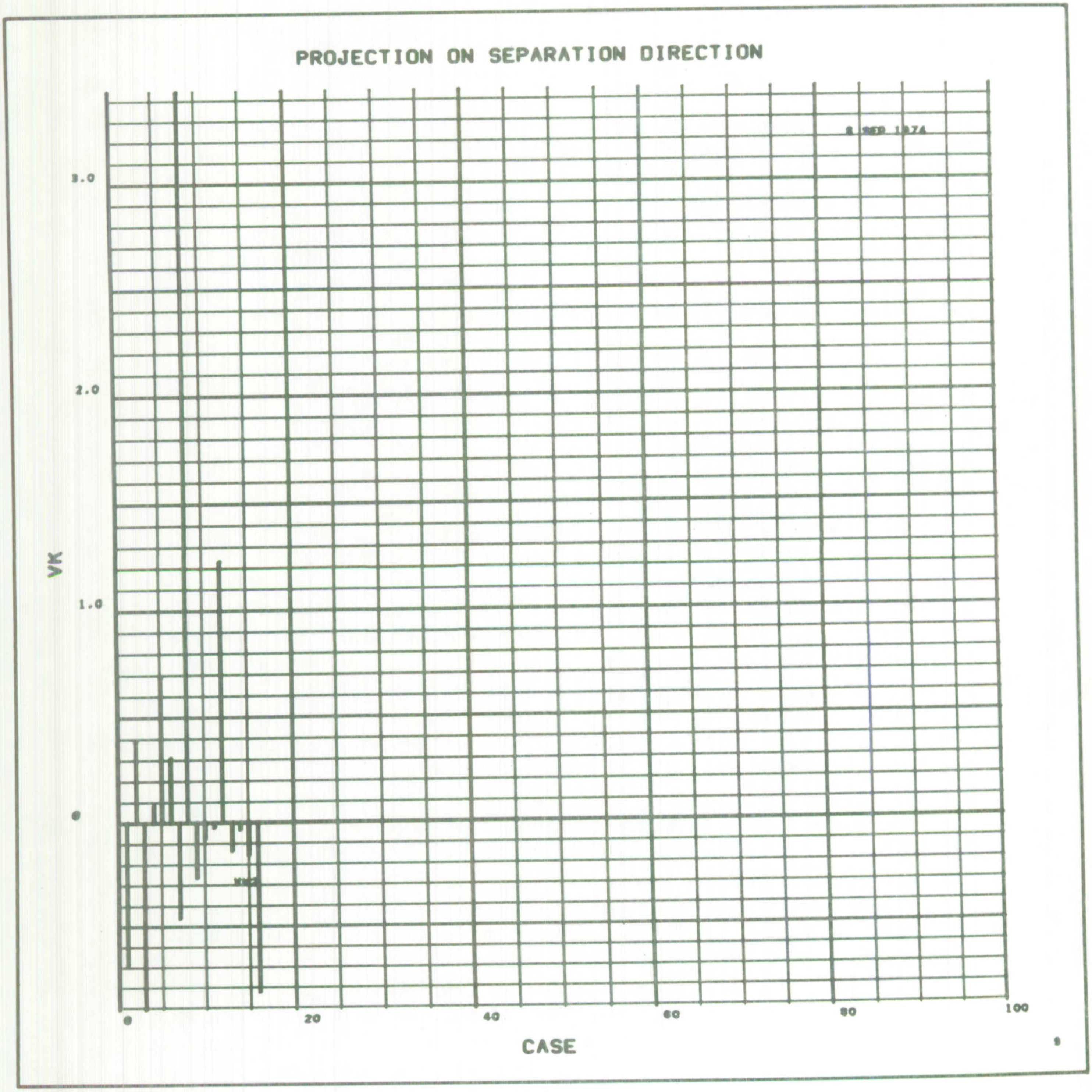


FIGURE 4.28B - BAR CHART SHOWING THE MAGNITUDE OF PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE LOW VARIATION CASES ON THE REFERENCE BASE-101 TARGETS

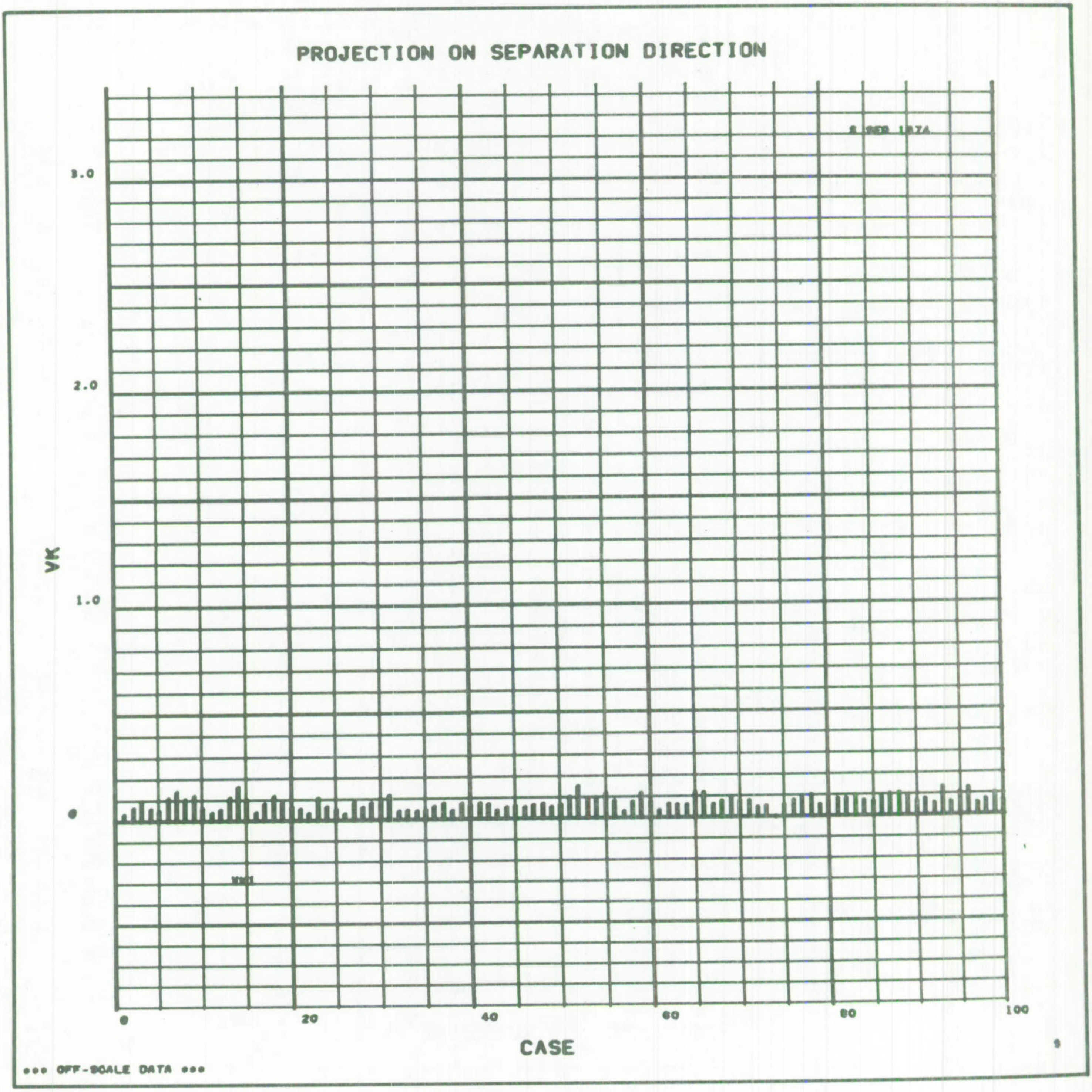


FIGURE 4.29 - RELATIVE IMPORTANCE OF EACH OPTIMAL COORDINATE TO MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION DEVELOPED USING THE LOW VARIATION CASES ON THE REFERENCE BASE

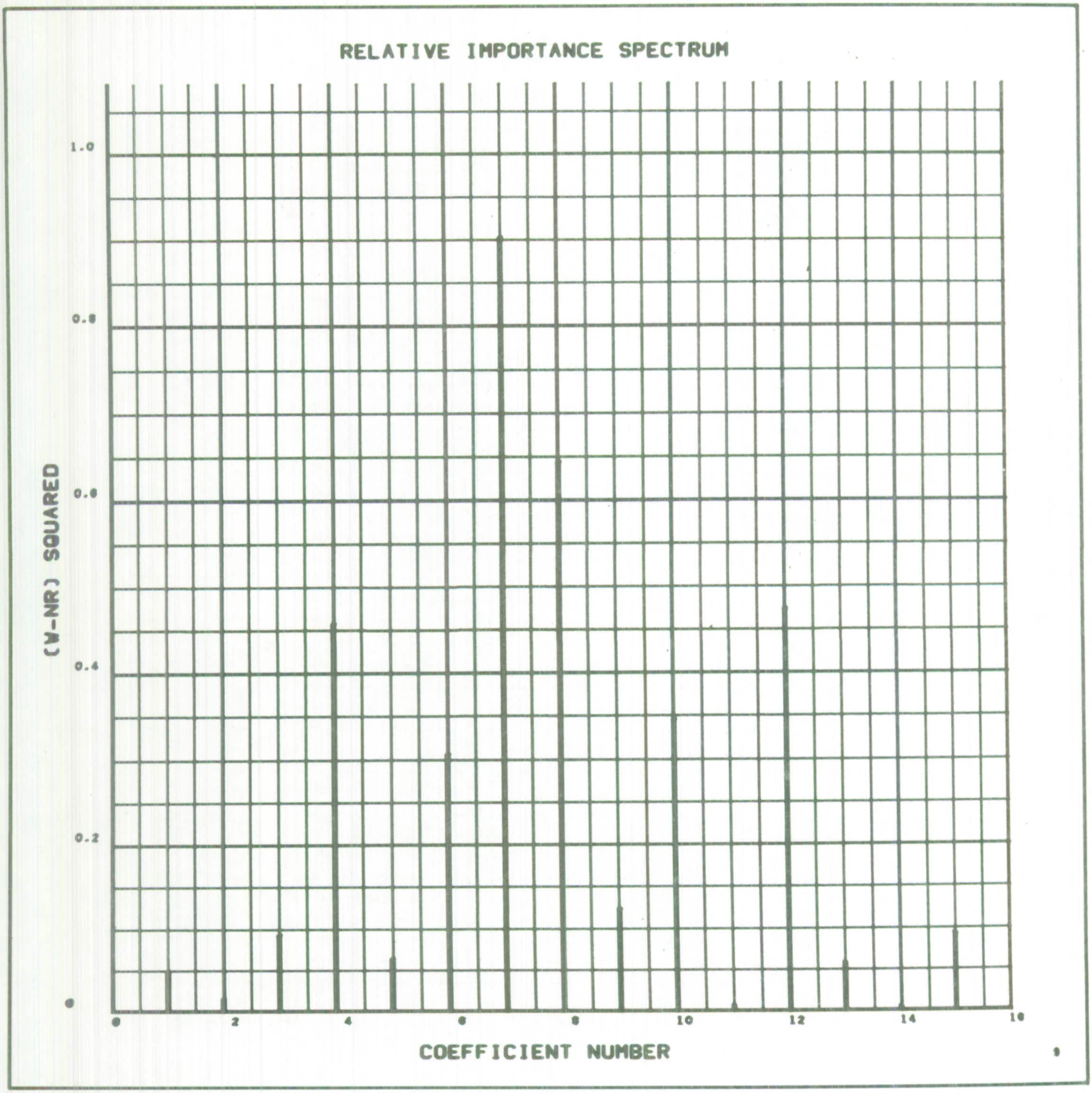


FIGURE 4.30 - RELATIVE IMPORTANCE OF SIGNAL ELEMENT CORRESPONDING TO INDEXING VARIABLE FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED USING THE LOW VARIATION CASES ON THE REFERENCE BASE

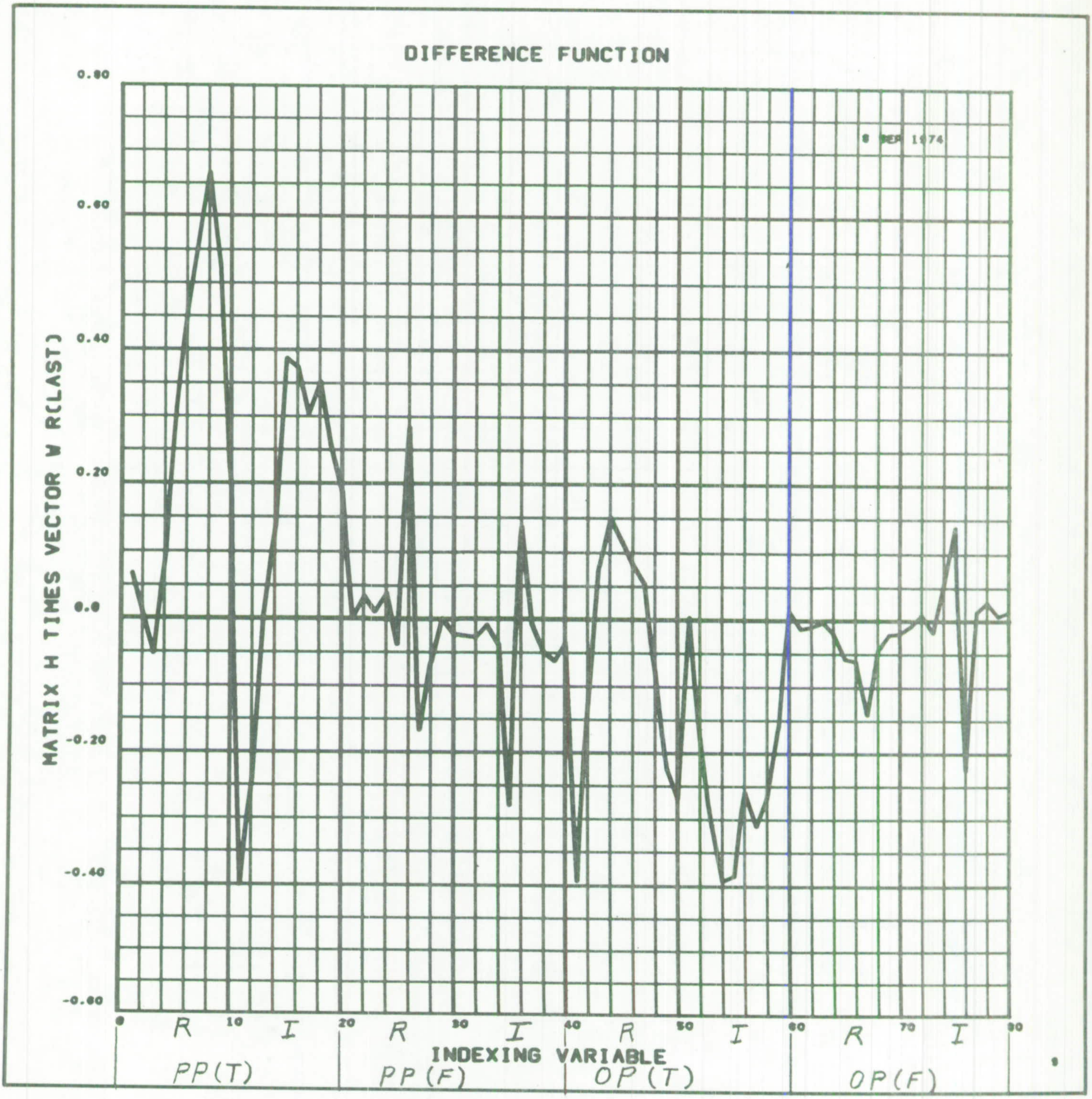


FIGURE 4.31 - CLASSIFICATION TRADE-OFF CURVES FOR PROJECTION OF TWO TARGETS ON THE INDEPENDENT CLASSIFIER MULTI-STEP CLASSIFICATION DIRECTION USING THE MINIMUM VARIATION CLASSIFIERS DEVELOPED ON REFERENCE BASE WITH BOTH HIGH AND LOW VARIATION CASES

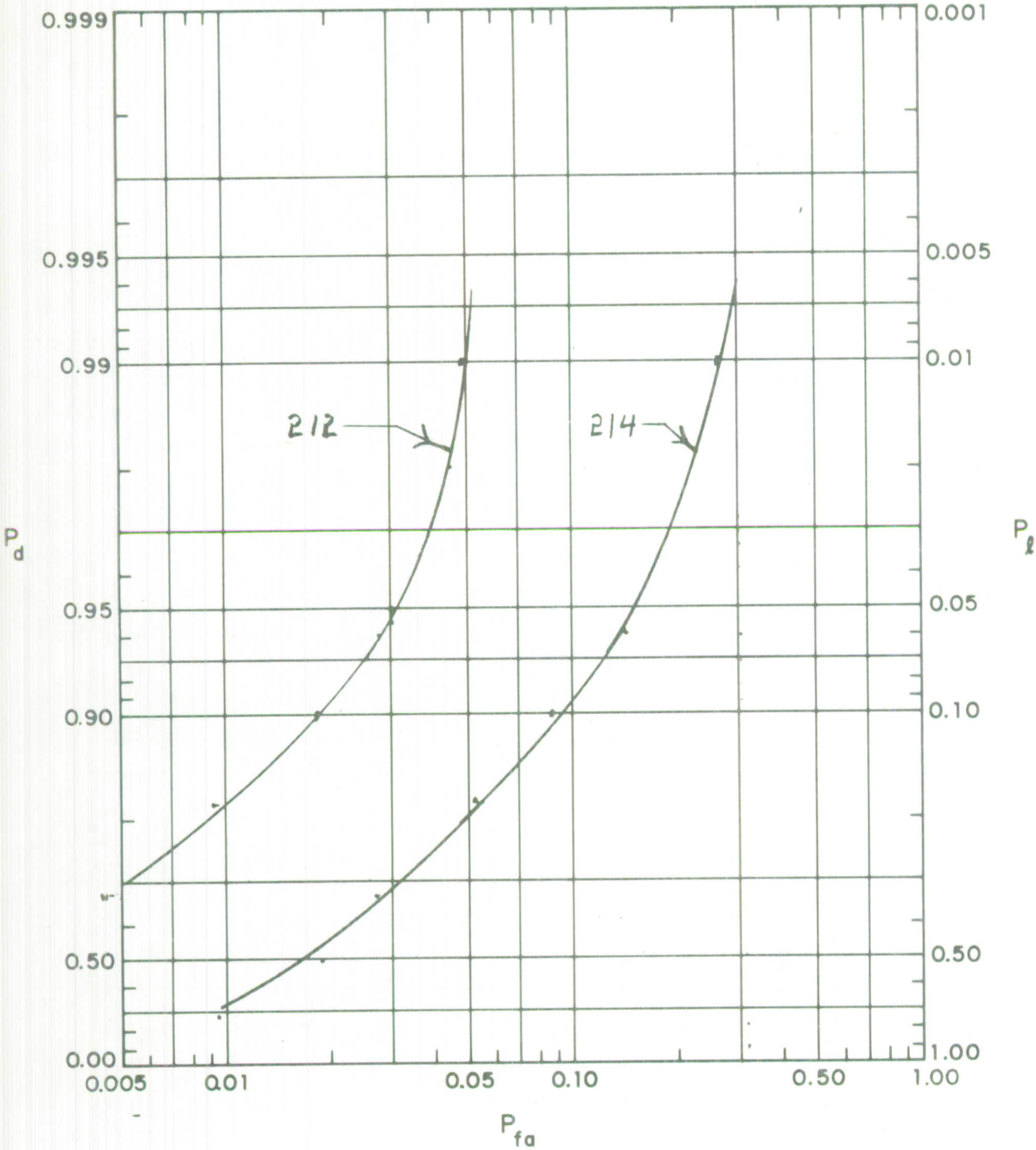


FIGURE 4.32A- BAR CHART SHOWING THE MAGNITUDE OF PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION DEVELOPED USING THE LOW VARIATION BASE - 21X TARGETS

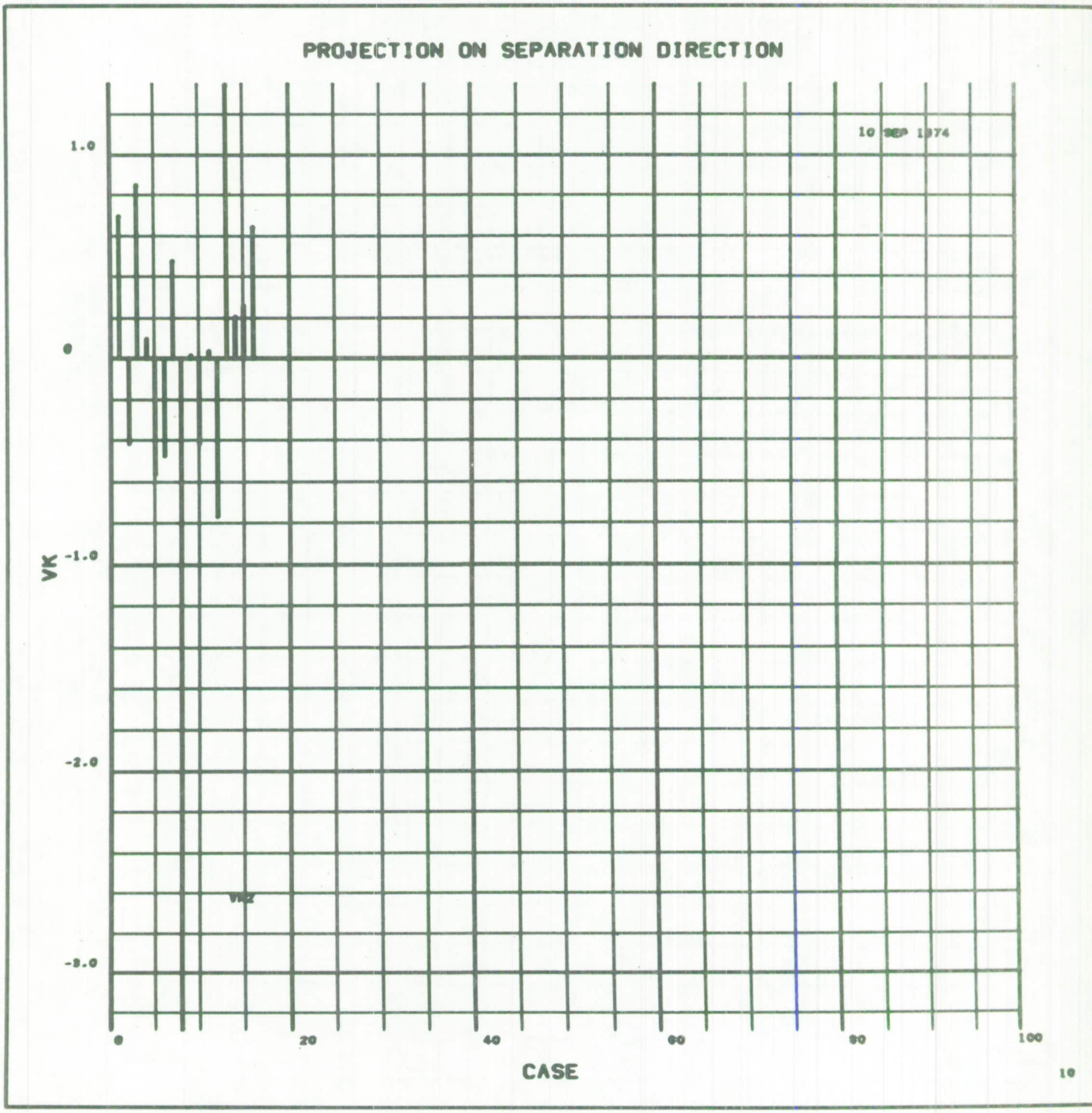
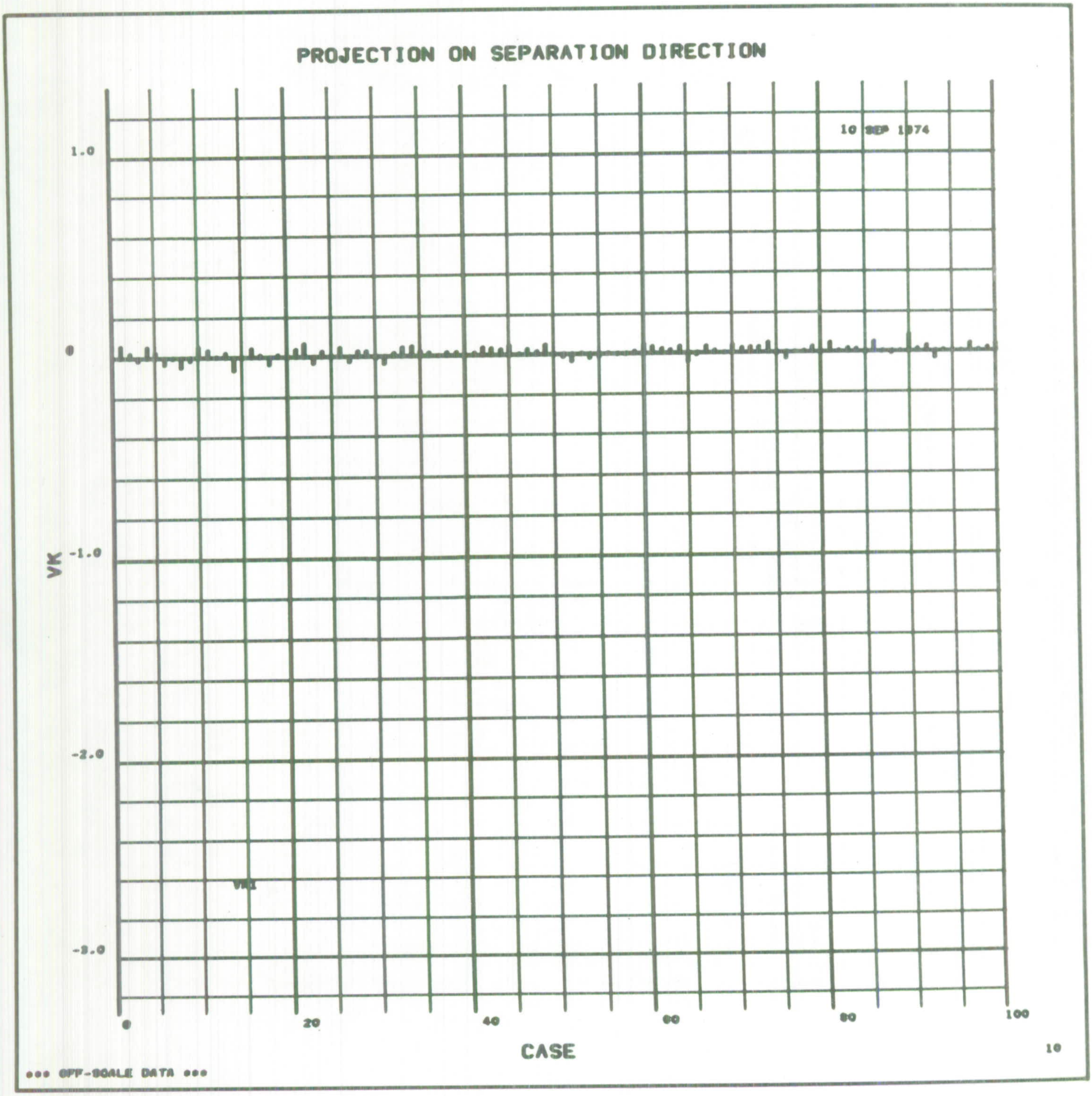


FIGURE 4.32B - BAR CHART SHOWING THE MAGNITUDE OF PROJECTION OF EACH LEARNING CASE ON THE MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION DEVELOPED USING THE LOW VARIATION BASE -101 TARGETS



procedure was not investigated further. Figures 4.33 and 4.34 give the relative importance spectrum and relative importance vector for the algorithm used to calculate the results presented in Figure 4.32. The spectra show that the second optimal direction of the low variation base dominates the minimum variation ratio classifier. Again, the relative importance vector presented in Figure 4.34 shows that this classifier is different from either the minimum variation ratio classifier developed on the reference base or the minimum variation ratio classifier developed on the reference base using only the 101 and 16-21X cases identified as being similar to the 101 Class through the scatter plot analysis.

FIGURE 4.33 - RELATIVE IMPORTANCE OF EACH OPTIMAL COORDINATE TO MINIMUM VARIATION RATIO CLASSIFICATION DIRECTION DEVELOPED ON LOW VARIATION BASE

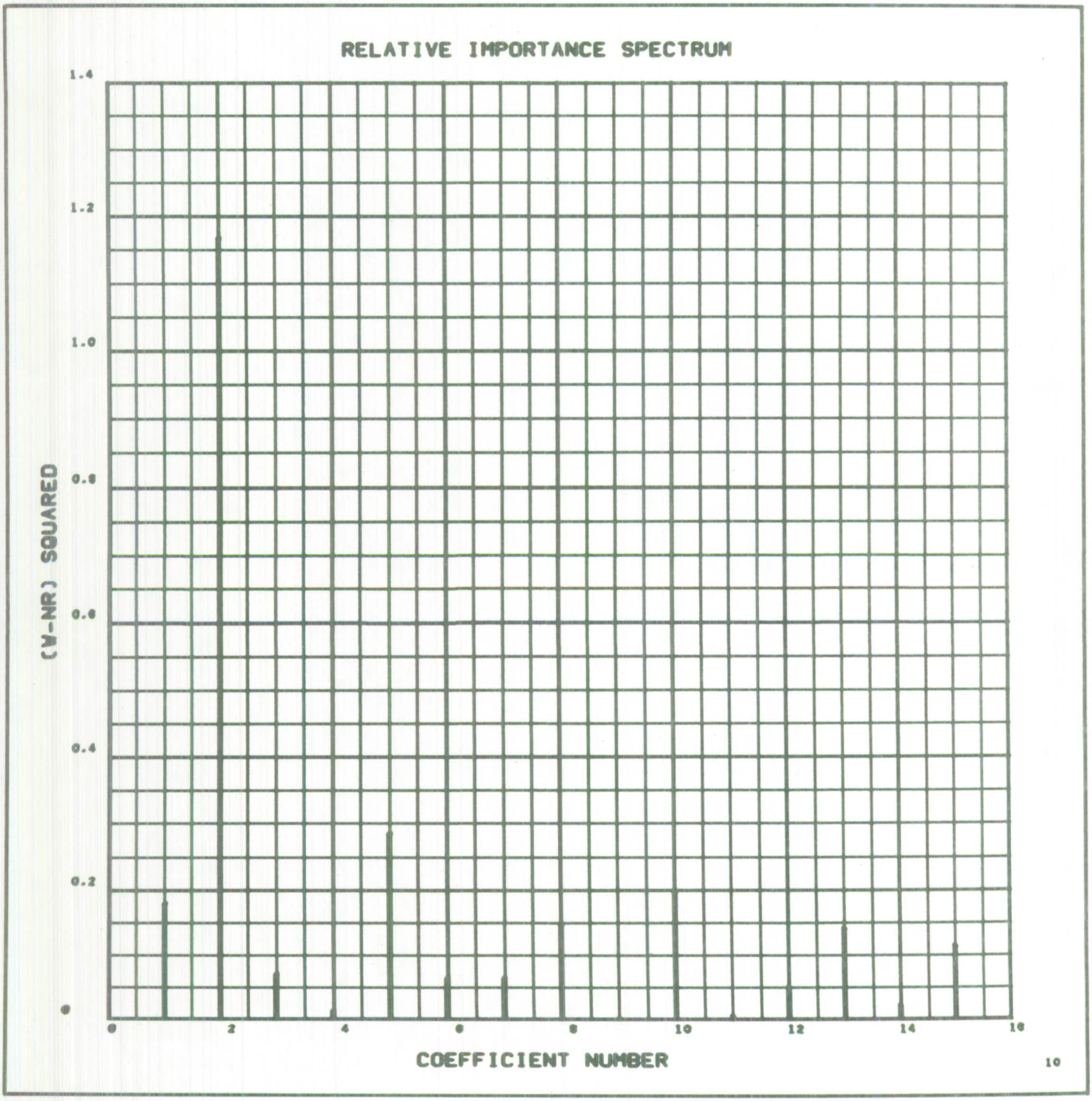
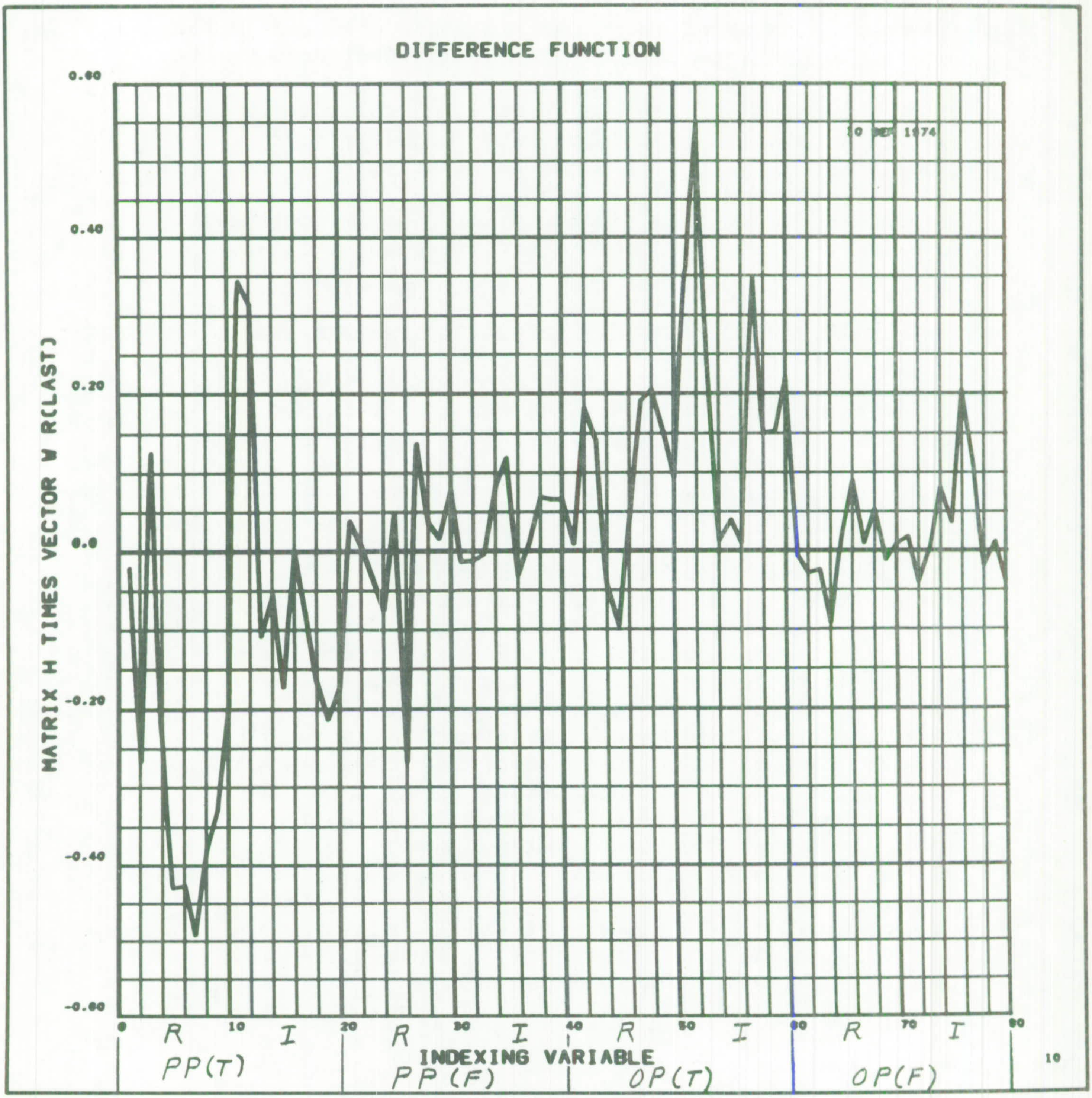


FIGURE 4.34 - RELATIVE IMPORTANCE OF SIGNAL ELEMENT CORRESPONDING TO INDEXING VARIABLE FOR MINIMUM VARIATION RATIO CLASSIFIER DEVELOPED ON LOW VARIATION BASE



5.0 ANALYSIS OF NON-LINEAR CLASSIFICATION SCHEMES

In this report, non-linear classification schemes are defined as any classification schemes for which the detection statistic is obtained by a non-linear calculation. With this definition all non trivial nearest neighbor calculations are considered non-linear since the distance calculation even in Euclidean space requires a non-linear operation. However, non-linear pre-processing of the data prior to calculation of the detection statistic does not make the classifier a non-linear classifier. Since in general, non-linear classifiers require considerable more computation their performance must show a corresponding improvement over the linear classifier before they can be justified.

5.1 Selection of Candidate Non-Linear Classifiers

Selection of the candidate non-linear classifiers to be considered in this study was constrained to include those classifiers for which the pertinent calculations already existed in the ADAPT family of programs. These classifiers may be visualized as belonging to two general groups. The first are the non-linear classifiers based on the Euclidean metric and the second group are the non-linear classifiers based on the Gaussian metric or the Mahalanobis distance. In either case, there are several options available for using the detection statistic in a multi-class problem. One group of schemes for utilizing the detection statistic might be called the nearest centroid or selection of the most likely class. Geometrically, this may be visualized as assigning each case to that class whose centroid is closest to the case. Although there are five individual classes for the present problem, there is only one decision of interest; namely, whether the target belongs to the 101 class or any one of the 21X Classes. This introduces two options into the nearest centroid approach. One may either use the centroid of all members of the 21X Class in the analysis or one may perform the analysis for each of the 211 through 214 Classes independently.

An alternate approach to selecting the most likely case or the nearest centroid, is to threshold the distance of each case from the class centroid. These detection statistics may be thresholded in the same manner as was done for the linear discriminant. In the present problem, this introduces up to six potential detection statistics. Thus, we have a total of eight candidate classification schemes available. These eight schemes will be evaluated against the reference, the principal polarization time, the square, the 3-Class II target and the Class II only bases. This gives a potential set of 40 candidate classification schemes for the Euclidean and maximum likelihood detection statistics.

The Euclidean detection statistic that will be used for this study is the simple Euclidean distance between the centroid of the class and each of the individual cases. The Gaussian statistic used in this analysis is proportional to the probability that the case belongs to Class ℓ based on the assumption that the cases have Gaussian distributions. The specific Gaussian detection statistic used in the present analysis is given by:

$$DK_{\ell} = -\tfrac{1}{2} \ln(|G_{\ell}|) - \tfrac{1}{2} (Y-\bar{Y}_{\ell})^{t} G_{\ell} (Y-\bar{Y}_{\ell})$$

where G_{ℓ} is the inverse of the covariance matrix for Class ,
$\bar{Y}_{\ell}$ is the class mean, and Y_{ℓ} is the ADAPT coefficient associated with the data history.

5.2 Performance Comparison of Candidate Non-Linear Classifiers Using Different ADAPT Bases

Although there are a total of 40 candidate classification schemes available for the Euclidean and Gaussian metric detection statistics, the use of the centroid of the entire 21X Class will only be evaluated on the principal polarization time base. Thus, the total number of candidates evaluated for each metric is reduced to 32. The initial gross comparison of these 64 classification schemes is summarized in Table 5.1.

This table gives the total number of errors which are made in the set of 240 learning cases and 360 test cases. The thresholded detection statistics are evaluated for the special case where the threshold is selected so that there are no errors in the classification of the 101 Class and the errors counted are the errors which are observed in the 21X Classes. The training data was used to develop the ADAPT optimal representation for the nearest neighbor classifications using Euclidean detection statistic. No additional training is required. For the Gaussian metric detection statistic, the training data was used to define the ADAPT representation and estimate the mean and standard deviations of the distribution function.

Table 5.1 illustrates many general conclusions regarding the value of the non-linear classifiers and the effects of the different bases used. For almost all cases except the square base the most effective classifier is the thresholding of either the Euclidean or the Gaussian detection statistic for the distance from the unknown case to the centroid of Class 101. For the Euclidean distance measure, the next best classifier is the distance related to the centroid of Class 212. For the Gaussian metric statistic, the distance to the nearest centroid provides the second best classifier. The number of test cases evaluated is insufficient to decide between the use of the Euclidean or Gaussian metrics for the performance of the thresholded detection statistic based on the centroid of the 101 Class.

Table 5.1 shows that the classification derived using the square base performed much less satisfactorily then any of the other bases evaluated. This is in sharp contrast to the results observed for the linear classifier which was presented in Section 4. However, this result was universal for all of the non-linear classification schemes considered and may be a consequence of the square pre-processing making the classifiers used depend on the fourth moments of the original data. Thus, all further evaluation of the non-linear classifier was restricted to bases developed without the square pre-processing. This also suggests that the use of the amplitude and phase form of the data may reduce the effectiveness of these non-linear classification schemes.

TABLE 5.1

PERFORMANCE SUMMARY TABLE

CLASSIFICATION SCHEME / BASE	TOTAL NUMBER OF CASES MISSED OUT OF 600 CASES TESTED						
		THRESHOLDING OF DETECTION STATISTIC = DISTANCE TO CENTROID OF:					
	NEAREST CENTROID	CLASS 101	CLASS 21X	CLASS 211	CLASS 212	CLASS 213	CLASS 214
REFERENCE							
EUCLIDEAN	163	0	-	37	1	40	19
GAUSS METRIC	5	0	-	13	22	70	6
PP-TIME ONLY							
EUCLIDEAN ()= 2-Class	149 (269)	4 (4)	(25)	43	6	46	29
GAUSS METRIC ()= 2-Class	8 (3)	1 (1)	(59)	76	107	106	58
SQUARE (10-DIM)							
EUCLIDEAN	211	272	-	340	273	307	271
GAUSS METRIC	165	250	-	286	278	284	271
3-21X							
EUCLIDEAN	180	0	-	37	1	39	-
GAUSS METRIC	4	0	-	42	16	62	-
21-X-ONLY							
EUCLIDEAN	160	0	-	37	1	40	19
GAUSS METRIC	4	0	-	13	23	68	10

The selection of the most likely class
on the Gaussian statistics performed significantly better than the selection of the most likely case based on the Euclidean distance of the centroids of each of the classes. These classification schemes were generally evaluated based on assigning the target to one of the five classes, however errors between the members of the 21X Class were not considered. If the nearest class was the 21X Class and the target was any one of the 21X targets, it was considered to be correctly classified. The use of only two classes in this procedure was also investigated using the principal polarization time only base. For this investigation, a separate classification rule was made based on the distance to the centroid of the entire 21X Class. The results based on this procedure are indicated in Table 5.1 by the values enclosed in parentheses. This procedure for the nearest centroid classifications showed no improvement in the case of the Euclidean metric and a reduction in the number of errors from eight to three in the case of the Gaussian metric. However, for the reference base, this method must yield at least five errors. This is because classification performance on the 101 Class is identical regardless of whether the entire 21X Class is used to compute the classification scheme or each of the 21X Classes is calculated separately. Thus, the number of errors occurring in the 101 Class will remain the same. Since all five of the errors shown on Table 5.1 for the Gaussian metric nearest centroid scheme for the reference base occurred in the 101 Class, these same five errors will occur if only two classes are used in the analysis.

With the exception of the significant degradation of the performance when the square base was used the effect of varying the base used was essentially as would be expected

from the results presented in Sections 3 and 4. The classifications based on the principal polarization time only base were slightly inferior to the corresponding classifications using the reference base. The classification using the 21X only base was very similar to using the reference base and the classification performance using the 3-Class II base was slightly less like the performance using the reference classification. These results were true in general regardless of which of the classification schemes shown on Table 5.1 were utilized.

Based on the results presented in Section 3, no differences were expected between the results obtained using the Class II only base and the results obtained using the reference base. Thus, the discrepancies in performance indicated by Table 5.1 between these two bases were investigated in detail to determine the extent to which the differences occur. One significant and unexplained difference is the performance of the nearest centroid classification scheme using Gaussian metric on the first case. This difference indicates that the representation of this first case on the two bases is quite different. Examination of the overall summary statistics and a review of a large number of the individual cases involved showed that the first case was the only case examined for which such large differences occurred.

Except for the first case the differences in performance shown on Table 5.1 using the Class II only and the reference base are due to small differences in the detection statistics. Table 5.1 shows that on the reference base there are 163 errors while on the Class II only base there are 160 errors. The three errors occurring on the reference base which did not occur on the Class II only base occurred in Cases 193, 261, and 292. These differences were due to very small differences in the detection statistic for all three of these cases. For example, for Case No. 193, the distance to the Class I centroid was .01943 for the reference base and .01949 for the Class II only base and the corresponding distances to the centroid of the third class were .019586

for the reference base and .019403 for Class II only base. Thus, we see that although these values agree to almost three significant figures the order and therefore the classification results have changed. Thus, we conclude that in general with the exception of the square pre-processing and the first case in the 600 case test set the effect of the base modifications on the non-linear classifiers are the same as was observed for the effect of these base modifications on the performance of the linear classifiers presented in Section 4.

Table 5.1 suggests that the best non-linear detection statistics for thresholding considered in this study is either the Gaussian or Euclidean statistic associated with Class 101. Thus, we shall consider the performance of these detection statistics for the reference, principal polarization time only, 3-Class II and Class II only bases in more detail. Referring to Table 5.1, we see that there were no errors in the classification as evaluated against the 600 case data set for the reference, the 3-Class II, and the Class II only bases. There is one error for the principal polarization time only base using the Gaussian metric detection statistic and this occurred for one of the test cases associated with the 214 Class. There were four errors in the classification based on the Euclidean distance to the Class 101 centroid. Two of these errors occurred in the 212 Class one each in the learning and test data and the other two errors occurred in the learning data for the 214 Class. Although the number of cases involved is not sufficient to arrive at high confidence conclusions, the trend indicated by these errors is similar to the trend which was observed for the effects of the various target types on classification performance for the linear classifiers presented in Section 4.

To provide a better measure of performance based on the Class 101 detection statistics, statistical summaries are presented for these detection statistics in Tables 5.2 through 5.9. These tables have the same format and present the same information regarding the Euclidean or Gaussian detection statistics based on the centroid of Class 101 that was presented for the linear classifier detection statistics in Tables 4.1 through 4.10. Specifically, these tables present a class identification where Classes 1 and 10 are the 101 test and learning classes, respectively,

TABLE 5.2 - DETECTION SUMMARY-EUCLIDEAN DISTANCE TO CENTROID OF CLASS 101 DEVELOPED FROM THE REFERENCE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID 5.2
1	100	0.4290E 02	0.9294E 01	0.5215E 02	73	-0.1172E 02	28
2	40	-0.1041E 06	0.2013E 06	-0.1175E 04	13	-0.1012E 07	25
3	60	-0.4755E 05	0.8120E 05	-0.1813E 04	20	-0.4042E 06	9
4	40	-0.2315E 04	0.4167E 04	-0.1843E 03	11	-0.2469E 05	27
5	60	-0.4306E 04	0.9655E 04	-0.1526E 03	45	-0.6895E 05	9
6	40	-0.3597E 05	0.4264E 05	-0.9268E 03	31	-0.2129E 06	30
7	60	-0.3937E 05	0.4741E 05	-0.8741E 03	12	-0.2893E 06	40
8	40	-0.6625E 04	0.1679E 05	-0.3879E 02	36	-0.9741E 05	22
9	60	-0.5141E 04	0.1450E 05	-0.3319E 02	9	-0.1033E 06	41
10	100	0.4581E 02	0.3621E 01	0.5163E 02	67	0.2597E 02	37

TABLE 5.3 - DETECTION STATISTIC SUMMARY-GAUSSIAN METRIC OF CLASS 101 DEVELOPED FROM REFERENCE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID 5.3
1	100	0.1299E-01	0.5717E-02	0.3749E-01	28	0.2598E-02	61
2	40	0.3718E 02	0.5341E 02	0.2241E 03	2	0.3690E 00	13
3	60	0.2068E 02	0.2719E 02	0.1365E 03	9	0.7770E 00	23
4	40	0.1077E 01	0.2359E 01	0.1485E 02	27	0.8798E-01	11
5	60	0.2410E 01	0.5620E 01	0.2838E 02	14	0.1179E 00	48
6	40	0.1465E 02	0.1513E 02	0.5352E 02	20	0.7741E 00	31
7	60	0.1549E 02	0.1684E 02	0.8382E 02	40	0.3510E 00	18
8	40	0.4065E 01	0.1132E 02	0.6750E 02	22	0.4096E-01	37
9	60	0.2416E 01	0.6008E 01	0.3831E 02	41	0.5313E-01	9
10	100	0.1426E-01	0.5140E-02	0.2738E-01	36	0.3421E-02	100

TABLE 5.4 - DETECTION STATISTIC SUMMARY-EUCLIDEAN DISTANCE TO CENTROID OF CLASS 101 DEVELOPED FROM PRINCIPAL POLARIZATION TIME BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID 5.4
1	100	0.4090E 02	0.1070E 02	0.4838E 02	94	-0.9277E 01	62
2	40	-0.6017E 05	0.9782E 05	-0.2461E 03	13	-0.4458E 06	2
3	60	-0.4074E 05	0.5371E 05	-0.1248E 04	20	-0.2554E 06	26
4	40	-0.2218E 04	0.6061E 04	-0.6391E 02	8	-0.3884E 05	27
5	60	-0.4439E 04	0.1134E 05	-0.3299E 02	23	-0.8081E 05	9
6	40	-0.2049E 05	0.2314E 05	-0.8544E 03	31	-0.9483E 05	8
7	60	-0.3085E 05	0.5089E 05	-0.4236E 03	60	-0.3432E 06	40
8	40	-0.5858E 04	0.1736E 05	-0.1211E 02	4	-0.1025E 06	22
9	60	-0.4924E 04	0.1844E 05	0.1629E 02	17	-0.1408E 06	41
10	100	0.4395E 02	0.4122E 01	0.4850E 02	87	0.2124E 02	37

TABLE 5.5 - DETECTION STATISTIC SUMMARY-GAUSSIAN METRIC OF CLASS 101 DEVELOPED FROM PRINCIPAL POLARIZATION TIME BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID 5.5
1	100	0.1632E-01	0.5342E-02	0.3636E-01	28	0.4941E-02	67
2	40	0.2367E 02	0.3093E 02	0.1609E 03	2	0.8666E-01	13
3	60	0.1773E 02	0.1644E 02	0.8569E 02	26	0.7356E 00	23
4	40	0.1010E 01	0.2591E 01	0.1658E 02	27	0.3149E-01	8
5	60	0.2301E 01	0.5728E 01	0.2861E 02	14	0.3081E-01	39
6	40	0.9544E 01	0.1102E 02	0.4930E 02	20	0.6192E 00	16
7	60	0.1244E 02	0.1660E 02	0.8130E 02	40	0.3052E 00	18
8	40	0.3620E 01	0.1069E 02	0.6428E 02	22	0.2286E-01	17
9	60	0.2265E 01	0.6790E 01	0.4760E 02	41	0.3872E-01	17
10	100	0.1679E-01	0.4649E-02	0.2939E-01	51	0.4899E-02	100

TABLE 5.6 - DETECTION STATISTIC SUMMARY-EUCLIDEAN DISTANCE TO CENTROID OF CLASS 101 DEVELOPED FROM THREE CLASS II BASE

5.6

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.4324E 02	0.8716E 01	0.5246E 02	73	0.2705E 01	28
2	40	-0.1023E 06	0.1951E 06	-0.1195E 04	13	-0.9645E 06	25
3	60	-0.4784E 05	0.8099E 05	-0.1870E 04	20	-0.3723E 06	9
4	40	-0.2259E 04	0.3765E 04	-0.1635E 03	11	-0.2151E 05	27
5	60	-0.4365E 04	0.1014E 05	-0.1767E 03	37	-0.7285E 05	9
6	40	-0.3787E 05	0.4454E 05	-0.8557E 03	31	-0.2214E 06	30
7	60	-0.4033E 05	0.4887E 05	-0.8670E 03	12	-0.3018E 06	40
8	40	-0.5353E 04	0.1266E 05	-0.4077E 02	37	-0.7356E 05	22
9	60	-0.4956E 04	0.1304E 05	-0.2930E 02	7	-0.8997E 05	41
10	100	0.4620E 02	0.3473E 01	0.5200E 02	67	0.2826E 02	37

TABLE 5.7 - DETECTION STATISTIC SUMMARY- GAUSSIAN METRIC OF CLASS 101 DEVELOPED FROM THREE CLASS II BASE

5.7

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.1237E-01	0.5747E-02	0.3593E-01	28	0.3232E-02	73
2	40	0.3737E 02	0.5357E 02	0.2244E 03	2	0.3740E 00	13
3	60	0.2010E 02	0.2686E 02	0.1298E 03	9	0.7599E 00	23
4	40	0.1057E 01	0.2209E 01	0.1381E 02	27	0.8508E-01	11
5	60	0.2341E 01	0.5327E 01	0.2682E 02	14	0.1236E 00	48
6	40	0.1489E 02	0.1526E 02	0.5220E 02	20	0.7603E 00	31
7	60	0.1545E 02	0.1630E 02	0.7952E 02	40	0.3581E 00	18
8	40	0.3185E 01	0.8813E 01	0.5362E 02	22	0.3936E-01	37
9	60	0.2294E 01	0.5571E 01	0.3486E 02	41	0.5173E-01	7
10	100	0.1370E-01	0.5479E-02	0.2723E-01	36	0.2856E-02	67

TABLE 5.8 - DETECTION STATISTIC SUMMARY-EUCLIDEAN DISTANCE TO CENTROID OF CLASS 101 DEVELOPED FROM CLASS II ONLY BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.4290E 02	0.9325E 01	0.5217E 02	73	-0.1194E 02	28
2	40	-0.1041E 06	0.2013E 06	-0.1176E 04	13	-0.1012E 07	25
3	60	-0.4757E 05	0.8128E 05	-0.1815E 04	20	-0.4044E 06	9
4	40	-0.2313E 04	0.4156E 04	-0.1844E 03	11	-0.2461E 05	27
5	60	-0.4306E 04	0.9656E 04	-0.1523E 03	45	-0.6898E 05	9
6	40	-0.3601E 05	0.4265E 05	-0.9286E 03	31	-0.2127E 06	30
7	60	-0.3945E 05	0.4747E 05	-0.8770E 03	12	-0.2896E 06	40
8	40	-0.6628E 04	0.1680E 05	-0.3813E 02	36	-0.9744E 05	22
9	60	-0.5128E 04	0.1441E 05	-0.3341E 02	9	-0.1027E 06	41
10	100	0.4583E 02	0.3609E 01	0.5165E 02	67	0.2617E 02	37

TABLE 5.9 - DETECTION STATISTIC SUMMARY- GAUSSIAN METRIC OF CLASS 101 DEVELOPED FROM CLASS II ONLY BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	100	0.1297E-01	0.5717E-02	0.3750E-01	28	0.2620E-02	61
2	40	0.3718E 02	0.5341E 02	0.2241E 03	2	0.3693E 00	13
3	60	0.2067E 02	0.2723E 02	0.1365E 03	9	0.7766E 00	23
4	40	0.1075E 01	0.2355E 01	0.1482E 02	27	0.8854E-01	11
5	60	0.2405E 01	0.5603E 01	0.2830E 02	14	0.1176E 00	48
6	40	0.1466E 02	0.1512E 02	0.5352E 02	20	0.7737E 00	31
7	60	0.1550E 02	0.1684E 02	0.8382E 02	40	0.3512E 00	18
8	40	0.4057E 01	0.1129E 02	0.6734E 02	22	0.4091E-01	37
9	60	0.2413E 01	0.5998E 01	0.3823E 02	41	0.5314E-01	9
10	100	0.1422E-01	0.5133E-02	0.2737E-01	36	0.3393E-02	100

the even Classes 2 through 8 represent the training data for Classes 211 through 214, respectively, and the odd Cases 3 through 9 represent the 340 independent test cases for Classes 211 through 214, respectively. The table presents the number of cases in each class, the mean and standard deviation of the detection statistic as well as the max and min value and identifies for which case the max and min value occur.

Examinations of Tables 5.2 through 5.9 provide additional verification of the general conclusions already obtained from the Class 101 column of Table 5.1 and in general extend these conclusions to the bases for which no conclusions could be reached regarding the performance of these classifiers.
For example, if one examines Table 5.2, we find that the spread between the minimum value of Class 101 (-11.72) and the maximum value associated with Classes 211 through 214 is the greatest for the 211 Class the smallest for the 214 Class and the second smallest for the 212 Class. Thus, the difficulty in classification of the various targets is very similar to that experienced for the linear classifiers. Similarly, comparing these results with those presented in Table 5.3 shows that the relative difficulty for classification of each of the 21X target types was the same based on the Euclidean detection statistic. The spread between Class 8 minimum value and the Class 1 maximum value is very small both in terms of the associated mean values and standard deviations. This suggests that the result observed for the principal polarization time only base, i.e. slightly improved classification based on the Gaussian statistic, is probably also true for the classification using the reference base. The degree of similarity between the Class II only base and the reference base can also be seen by comparison of Tables 5.2 and 5.3 with Tables 5.8 and 5.9. Comparison of these tables with Tables 5.6 and 5.7 shows that there is less similarity between the 3-Class II base and the reference base.

Thus, the results of Table 5.1 and Tables 5.2 through 5.9 suggest that the best performance using the non-linear classifier will be obtained using the reference base. The best classification procedure appears to use the Gaussian metric that the target belongs to Class 101 as a thresholded

detection statistic. The major limitation on these conclusions is related to the small number of independent test cases which were available in the 600 case test set. Although these cases appeared sufficient to reach the qualitative conclusions discussed to this point, they are insufficient to allow an estimate of the expected performance of the recommended classifier on the recommended base.

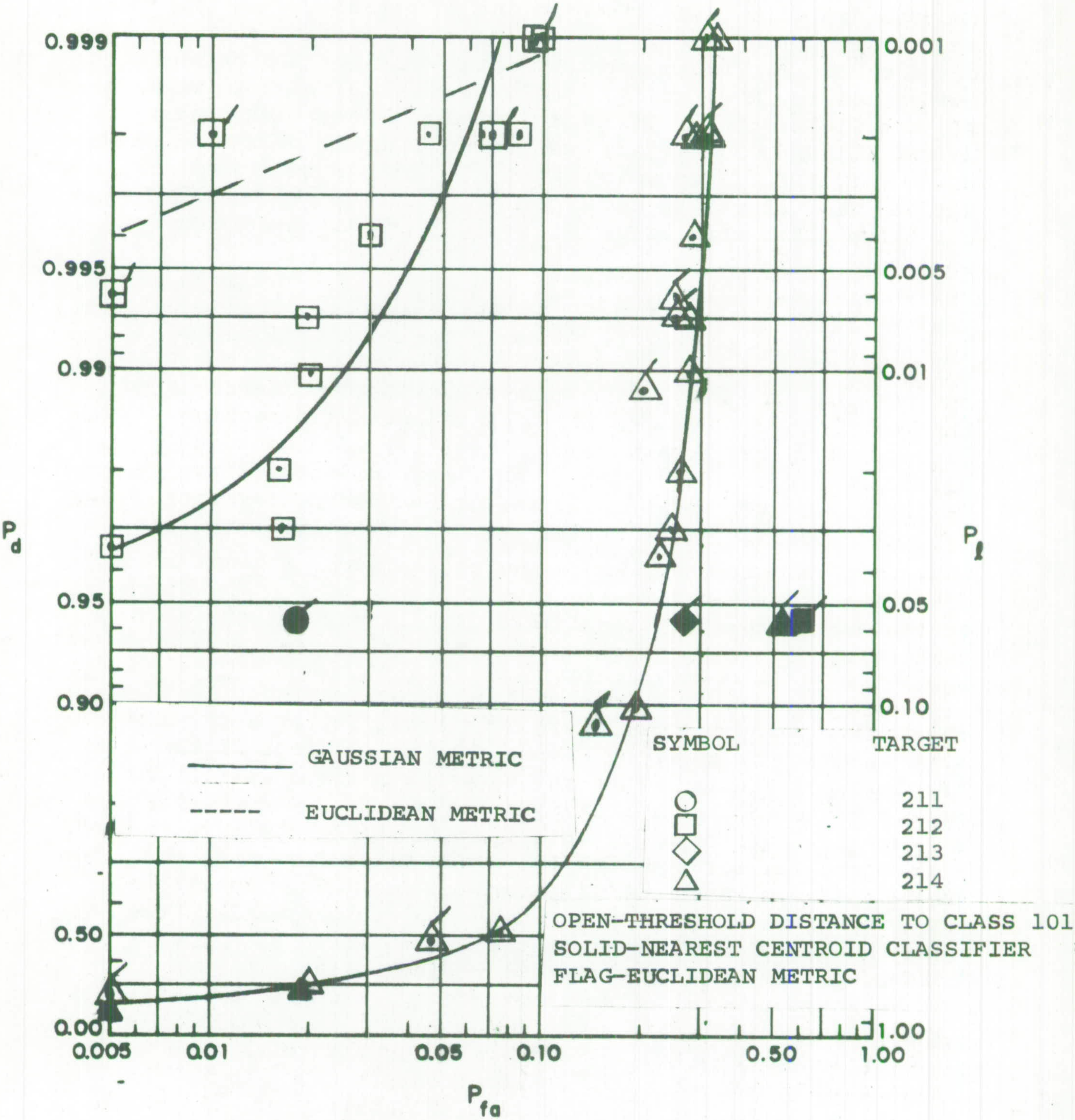
5.3 Performance of Non-Linear Classifiers on 1800 Case Test Set

Based on the results presented in Section 5.2, the proof testing of the non-linear classifiers was only performed on the classification schemes based on the detection statistics associated with the Class 101 centroid using the reference base. The performance achieved using these classification schemes should be the best performance possible from any of the non-linear classification schemes considered. Since the results of varying the base are either dramatic as in the case of the square base or consistent with the results expected from the analysis of the base in Section 3, further verification of the effect of modifying the base was not performed. The effect of varying the classification technique was checked using the reference base.

The performance of the classification based on the detection statistics associated with the centroid of Class 101 when applied to the 1800 independent test cases is summarized on Figure 5.1. The open symbols on this figure represent the experimental ROC data points from thresholding of the detection statistic associated with the centroid of the 101 Class. The solid symbols represent the results obtained from the nearest centroid classifier interpreted in terms of detection probability and false alarm rate. Those symbols which have been flagged are based on the Euclidean metric and the unflagged symbols are based on the Gaussian likelihood metric. The results for the thresholded classifiers have been fitted by solid lines for the Gaussian metric and dash lines for the Euclidean metric.

Examination of Figure 5.1 shows that the general results presented in Table 5.1 are confirmed. For the Euclidean metric, the thresholding classifier shows a better performance then the nearest centroid classifier. For the Gaussian likeli-

FIGURE 5.1 - CLASSIFICATION TRADE-OFF CURVES COMPARING PERFORMANCE OF THE NON-LINEAR CLASSIFICATION SCHEMES ON 1800 INDEPENDENT TEST CASES



OPEN - THRESHOLD DISTANCE TO CLASS 101
SOLID - NEAREST CENTROID CLASSIFIER
FLAG - EUCLIDEAN METRIC

—— GAUSSIAN METRIC
- - - EUCLIDEAN METRIC

5.1

hood metric, the performance of the nearest centroid and thresholding classifier are about equal. However, it is easier to select a point on the ROC curve at which the algorithm is to be applied for the thresholding classifier. Again, in agreement with the results of Table 5.1 the thresholded classifier based on the Gaussian metric performed better then the thresholded classifier based on the Euclidean metric. Note, that for the thresholded metric, ROC curves could only be obtained for the 212 and 214 targets. The performance of these classifiers on the 211 and 213 targets was better than could be evaluated using only 1800 independent samples. That is, for the 211 and 213 targets at detection probabilities up to .999 the false alarm rate is probably significantly less than .005.

Comparison of Figure 5.1 with the results presented in Section 4 shows that in general, the results obtained using the best non-linear classifier evaluated are similar but possibly slightly worse than the results obtained using the best linear (i.e. the first ADAPT optimal direction classifier on the square base) classifier for the most difficult target (i.e. 214). The linear classifier is definitely better for the 212 target and the performance of both classifiers was better than could be evaluated for the other two targets. Since the classifiers with non-linear detection statistics require significantly more computation to implement than the classifiers with linear detection statistics, it is recommended that the linear classifiers be used.

The false alarm rate for the nearest centroid classifier using the Euclidean metric was sufficiently low that it could be shown on Figure 5.1 for all four of the target types studied. When the Gaussian maximum likelihood metric was used, the resulting point on the ROC curve was at such low detection probability that all of the targets except the 214 target had false alarm rates less than could be evaluated with the 200 independent test cases available for each of the different 21X targets. This was also true when the classification procedure was modified such that a target was assigned to Class 101 whenever Class 101 was not selected by the nearest centroid approach as the least likely class. For this case, the detection probability is clearly larger and the false alarm rates should also be larger. However, we again see that only the 214 target had any false alarms for the detection probability of .315 which occurred when this ground rule was used.

Tables 5.10 through 5.19 summarize the statistical properties of the detection statistics obtained when the 1800 case test set is projected on the reference base. These tables are in the same general format as the tables 5.2 through 5.9. There are only five classes on Tables 5.10 through 5.19. The first class is the one thousand 101 Class targets in the 1800 case test set. Classes 2 through 5 are the 200 cases for each of Targets 211 through 214. For each of these five classes, these tables present the number of cases in the class, the mean and standard deviation of the detection statistic and the max and min value and its associated case number. Table 5.10 presents these results for the Gaussian metric detection statistic. This is the summary of the data from which the unflagged symbols on Figure 5.1 were calculated. Tables 5.11 through 5.14 presents similar information for the Gaussian metric classification statistic based on the centroids of Classes 211 through 214, respectively. Examination of these tables will verify the results obtained in the initial analysis presented in Table 5.1 that the detection statistic based on the centroid of the 101 Class is the best classification statistic to use.

Table 5.15 presents the same information for the detection statistic equal to the Euclidean distance from the 101 Target. This is a summary of the data which was used to calculate the information indicated by the flagged symbols on Figure 5.1. Similarly, Tables 5.16 through 5.19 present the summary of the detection statistic equal to the distance to the centroid of the 211 through the 214 Classes, respectively. Again, comparison of Tables 516 through 5.19 with Table 5.15 verifies the results presented in Table 5.1 that the distance to the centroid of the 101 Class is the best of these detection statistics.

TABLE 5.10 - SUMMARY OF 1800 TEST CASE VALUES OF GAUSSIAN METRIC OF CLASS 101 PROJECTED ON REFERENCE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID 5.10
1	1000	-0.3805E 02	0.5395E 02	0.4992E 02	537	-0.3237E 03	906
2	200	-0.7331E 05	0.1155E 06	-0.1264E 04	166	-0.6822E 06	88
3	200	-0.1272E 05	0.4381E 05	-0.1416E 03	88	-0.3995E 06	12
4	200	-0.4525E 05	0.5647E 05	-0.5545E 03	2	-0.3612E 06	176
5	200	-0.9847E 04	0.4740E 05	0.2491E 02	200	-0.4459E 06	56

TABLE 5.11 - SUMMARY OF 1800 TEST CASE VALUES OF GAUSSIAN METRIC OF CLASS 211 PROJECTED ON REFERENCE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID 5.11
1	1000	-0.8022E 01	0.2150E 00	-0.7953E 01	73	-0.8130E 01	906
2	200	-0.2701E 02	0.3828E 02	-0.7882E 01	150	-0.3213E 03	58
3	200	-0.1095E 02	0.9231E 01	-0.7928E 01	131	-0.7445E 02	59
4	200	-0.2268E 02	0.2572E 02	-0.7981E 01	2	-0.2345E 03	176
5	200	-0.9939E 01	0.9390E 01	-0.7941E 01	136	-0.1080E 03	4

TABLE 5.12 - SUMMARY OF 1800 TEST CASE VALUES OF GAUSSIAN METRIC OF CLASS 212 PROJECTED ON REFERENCE BASE

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID 5.12
1	1000	0.2097E 02	0.6450E 00	0.2150E 02	566	0.1800E 02	287
2	200	-0.4084E 03	0.7061E 03	0.1539E 02	166	-0.4576E 04	88
3	200	-0.6401E 02	0.3709E 03	0.2064E 02	88	-0.3809E 04	160
4	200	-0.2285E 03	0.3272E 03	C.1952E 02	2	-0.1604E 04	83
5	200	-0.2418E 02	0.2069E 03	0.2140E 02	110	-0.1920E 04	56

TABLE 5.13 - SUMMARY OF 1800 TEST CASE VALUES OF GAUSSIAN METRIC OF CLASS 213 PROJECTED ON REFERENCE BASE

5.13

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	1000	0.5360E 00	0.5410E-01	0.6584E 00	597	0.3359E 00	287
2	200	-0.2269E 02	0.3031E 02	0.2244E 00	150	-0.1798E 03	88
3	200	-0.4161E 01	0.1958E 02	0.7117E 00	137	-0.1637E 03	140
4	200	-0.1217E 02	0.1725E 02	0.4477E 00	16	-0.1366E 03	183
5	200	-0.2369E 01	0.1459E 02	0.6970E 00	196	-0.1436E 03	56

TABLE 5.14 - SUMMARY OF 1800 TEST CASE VALUES OF GAUSSIAN METRIC OF CLASS 214 PROJECTED ON REFERENCE BASE

5.14

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	1000	0.1319E 02	0.4745E 00	0.1393E 02	637	0.1065E 02	287
2	200	-0.4378E 03	0.7430E 03	0.9819E 01	91	-0.5680E 04	88
3	200	-0.6964E 02	0.2705E 03	0.1360E 02	165	-0.1693E 04	84
4	200	-0.2227E 03	0.3114E 03	0.1067E 02	171	-0.2115E 04	183
5	200	-0.3731E 02	0.2505E 03	0.1387E 02	12	-0.2655E 04	4

TABLE 5.15 - SUMMARY OF 1800 TEST CASE VALUES OF EUCLIDEAN DISTANCE TO CENTROID OF CLASS 101 PROJECTED ON REFERENCE BASE

5.15

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	1000	0.3352E-01	0.1570E-01	0.1271E 00	906	0.5776E-02	479
2	200	0.2762E 02	0.3834E 02	0.2305E 03	88	0.5320E 00	91
3	200	0.5283E 01	0.1722E 02	0.1240E 03	12	0.8795E-01	165
4	200	0.1605E 02	0.1898E 02	0.1047E 03	193	0.2047E 00	2
5	200	0.3532E 01	0.1669E 02	0.1516E 03	4	0.1889E-01	146

TABLE 5.16 - SUMMARY OF 1800 TEST CASE VALUES OF EUCLIDEAN DISTANCE TO CENTROID OF CLASS 211 PROJECTED ON REFERENCE BASE

5.16

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	1000	0.1322E 01	0.1455E 00	0.1915E 01	365	0.9880E 00	533
2	200	0.2879E 02	0.3850E 02	0.2401E 03	88	0.3727E 00	127
3	200	0.6580E 01	0.1681E 02	0.1133E 03	84	0.7321E 00	131
4	200	0.1776E 02	0.1987E 02	0.1075E 03	193	0.5545E 00	171
5	200	0.4779E 01	0.1679E 02	0.1542E 03	4	0.7807E 00	191

TABLE 5.17 - SUMMARY OF 1800 TEST CASE VALUES OF EUCLIDEAN DISTANCE TO CENTROID OF CLASS 212 PROJECTED ON REFERENCE BASE

5.17

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	1000	0.4247E-01	0.1570E-01	0.1508E 00	906	0.9662E-02	494
2	200	0.2763E 02	0.3838E 02	0.2311E 03	88	0.5136E 00	91
3	200	0.5298E 01	0.1722E 02	0.1236E 03	12	0.9018E-01	88
4	200	0.1609E 02	0.1897E 02	0.1043E 03	193	0.2100E 00	2
5	200	0.3534E 01	0.1666E 02	0.1514E 03	4	0.1894E-01	155

TABLE 5.18 - SUMMARY OF 1800 TEST CASE VALUES OF EUCLIDEAN DISTANCE TO CENTROID OF CLASS 213 PROJECTED ON REFERENCE BASE

5.18

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	1000	0.6081E 00	0.8407E-01	0.7961E 00	18	0.4212E 00	228
2	200	0.2820E 02	0.3871E 02	0.2280E 03	88	0.4272E 00	150
3	200	0.5788E 01	0.1710E 02	0.1227E 03	12	0.1557E 00	137
4	200	0.1592E 02	0.1853E 02	0.1006E 03	176	0.6443E 00	171
5	200	0.4057E 01	0.1737E 02	0.1634E 03	4	0.1630E 00	76

TABLE 5.19 - SUMMARY OF 1800 TEST CASE VALUES OF EUCLIDEAN TO CENTROID OF CLASS 214 PROJECTED ON REFERENCE BASE

5.19

CLASS	NO IN CLASS	MEAN	STD-DEV	WKMAX	MAX-ID	WKMIN	MIN-ID
1	1000	0.1775E 00	0.3833E-01	0.3789E 00	906	0.9541E-01	994
2	200	0.2768E 02	0.3821E 02	0.2337E 03	88	0.4825E 00	183
3	200	0.5460E 01	0.1719E 02	0.1201E 03	12	0.1232E 00	25
4	200	0.1633E 02	0.1920E 02	0.1069E 03	193	0.2574E 00	2
5	200	0.3656E 01	0.1663E 02	0.1525E 03	4	0.8940E-01	12

REFERENCES

1) Hunter, H.E., "Task-1 Final Report - Application of ADAPT to Quick Look Classification of Composite Radar Signatures", ADAPT 73-3, November 1973.

DEPARTMENT OF THE AIR FORCE
HEADQUARTERS ELECTRONIC SYSTEMS DIVISION (AFSC)
LAURENCE G. HANSCOM FIELD, BEDFORD, MASSACHUSETTS 01730

REPLY TO ATTN OF: TML (Lincoln Lab)

18 December 1974

SUBJECT: ESD-TR-74-345, dtd Nov 74, "Final Report Application of Adapt to Quick Look Classification of Composite Radar Signatures"

TO: DDC/Air Force Liaison Representative
Cameron Station
Alexandria, VA 22314

1. I certify that the subject TR has been reviewed and approved for public release by the controlling office and the information office in accordance with AFR 80-45/AFSC Sup 1. It may be made available or sold to the general public and foreign nationals.

2. Distribution statement A appears on the subject TR and the DD Form 1473 as required by AFRs 80-44 and 80-45.

FOR THE COMMANDER

EUGENE C. RAABE, Lt Colonel, USAF
Chief, Lincoln Laboratory Project Office

Atch
ESD-TR-74-345
(12 copies)

AD NUMBER	PAPER COPY PRICE	DATE
AD- A004206	$	1/24/75

1. REPORT IDENTIFYING INFORMATION

A. ORIGINATING AGENCY Lincoln Laboratory
Massachusetts Institute of Technology

B. REPORT TITLE AND/OR NUMBER ADAPT Corporation Final Report
Report No. 74-4, dated November 1974.

C. MONITOR REPORT NUMBER
ESD-TR-74-345

D. PREPARED UNDER CONTRACT NUMBER(S)
F19628-73-C-0002 Purchase Order No. C-563

2. DISTRIBUTION STATEMENT

"A"